A unified framework for spline estimators

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades

– Doctor rerum naturalium –

der Georg-August-Universität Göttingen

im Promotionsprogramm "Mathematik"

der Georg-August University School of Science (GAUSS)

vorgelegt von

Katsiaryna Schwarz

aus

Minsk

Göttingen, 2012

Betreuungsausschuss

Prof. Dr. Tatyana Krivobokova, Institut für Mathematische Stochastik

Prof. Dr. Martin Schlather, Institut für Mathematik, Universtität Mannheim

Prof. Dr. Robert Schaback, Institut für Numerische und Angewandte Mathematik

Mitglieder der Prüfungskommission

Referent: Prof. Dr. Tatyana Krivobokova, Institut für Mathematische Stochastik

Korreferent: Prof. Dr. Martin Schlather, Institut für Mathematik, Universtität Mannheim

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Robert Schaback, Institut für Numerische und Angewandte Mathematik

Jun.-Prof. Dr. Andrea Krajina, Institut für Mathematische Stochastik

Prof. Dr. Russell Luke, Institut für Numerische und Angewandte Mathematik

Prof. Dr. Preda Mihailescu, Mathematisches Institut

Tag der mündlichen Prüfung: 24.01.2013

To my husband Niko Schwarz and to Tatyana Krivobokova

Table of Contents

List of Figures

1 Introduction

This dissertation investigates some aspects of spline smoothing. There are three spline-based methods for the approximation of statistical data: smoothing splines, regression splines and penalized splines. In this dissertation, we investigate penalized splines and their connection to smoothing and regression splines. After presenting the theoretical background, we study: a unified framework, the local and the global asymptotic properties of penalized splines.
A detailed overview of spline estimators together with formulas, notations and terminology is given in Chapter 2. In the following three paragraphs, we give a brief description, research methods and practical usage for each spline estimator.
A *smoothing spline* estimator arises as the solution to a certain variation problem and is a spline with knots at the observation points. A penalty parameter controls the trade-off between the fidelity to the data and the smoothness of the estimator. Smoothing splines are an old and well-studied technique. The main tool for studying their asymptotic properties is the reproducing kernel Hilbert spaces framework. A practical flaw of smoothing splines is that their parameter dimension is high, i.e. the number of parameters to estimate equals the number of observations. Hence, smoothing splines can be computationally expensive for large sets of data or in high dimensions.
A *regression spline* estimator is obtained as the least-square projection onto a spline space with fewer knots than the number of observations. The smoothness and the goodness of the estimator is controlled by the number and the position of the knots. The asymptotics of regression splines is well-studied and is based on the use of the results of Barrow and Smith (1978) on the best L_2-approximation of a smooth function by a spline set. Unlike smoothing splines, regression splines have a low parameter dimension. Their drawback is that the choice of the number and of the position of the knots is crucial and this choice is a non-trivial optimization problem.

A *penalized spline* estimator combines the projection onto a low-dimensional spline space with the roughness penalty. A penalized spline estimator is a trade-off between a smoothing and a regression spline, where the last two can be considered as particular instances of penalized splines. If the number of knots equals the number of observations, then penalized and smoothing splines coincide. If the penalty parameter equals zero, then penalized and regression splines coincide. The asymptotic properties of penalized splines are not well-studied. Claeskens et al. (2009) found that the global asymptotic properties of penalized spline estimators depend on transfer parameter k_q. This parameter depends on the number of knots, the penalty parameter, and the degree of splines. If k_q is bounded, then the global asymptotics of penalized splines is similar to that of regression splines, if k_q grows with the number of observations then the global asymptotics of penalized splines is similar to that of smoothing splines. Penalized splines are the most widely used estimators, since they circumvent the disadvantages of smoothing and regression splines. The statistical properties of smoothing, regression and penalized splines are investigated by different methods. In this dissertation we study all three estimators simultaneously, pursuing three steps:

1. we construct the Demmler and Reinsch (1975) basis
2. we investigate the properties of the equivalent kernel
3. we study the local asymptotics using kernel regression methods.

To make use of Fourier techniques, we assume that knots and observations are equidistant, and the unknown regression function and spline estimators are periodic. Additionally, we follow the idea of Claeskens et al. (2009) and investigate the asymptotic behavior of splines depending on k_q, where the extreme values of k_q correspond to smoothing or to regression splines. Let us outline each step.

In step one, we obtain the Demmler-Reinsch basis of the corresponding spline space. By definition, the basis is orthonormal under an inner product that depends on the observations points. Additionally, derivatives of the basis functions are L_2-orthogonal. Using known exponential splines and certain polynomials constructed for our problem, we derive a closed-form expression for the Demmler-Reinsch basis in our special case of equidistant knots and periodic splines. With its help, we find the Fourier coefficients and study the global asymptotic properties of periodic spline estimators. Any spline estimator can be represented with the help of the Demmler-Reinsch basis as a weighted sum of observations. The explicit form of the Demmler-Reinsch basis, which is unknown in general, allows us to study the weight function of all periodic spline estimator.

In step two, we examine the weight function of periodic splines, known as the equivalent kernel. Making use of the explicit expression for the Demmler-Reinsch basis, we obtain a closed-form expression for the equivalent kernel. This expression depends on the roots of certain polynomials, which are, in general, difficult to find. However, for low degree splines these roots are known and the equivalent kernel is given explicitly. The formula we obtain is both general and precise. It shows the dependence of spline estimators on the position between the knots and is used for obtaining the decay rate of the equivalent kernel. Apart from the closed-form expression, we find the moments of the equivalent kernel that play a significant role in obtaining the local asymptotic properties of periodic spline estimators.

Finally, in step three, we study the local asymptotics of all periodic spline estimators together. We look at periodic spline estimators as at kernel estimators. A kernel estimator is determined by a kernel satisfying certain moment conditions and by a bandwidth that controls the smoothness of the estimator. The local asymptotic properties of the kernel estimator are known and depend on the bandwidth. From the previous step, the equivalent kernel of periodic splines satisfies the necessary moment conditions. Hence, the missing part is the bandwidth. We find a bandwidth that depends on k_q and is universal for all spline estimators. With its help, we obtain the pointwise asymptotic behavior of periodic splines using known results from kernel regression.

The remainder of this dissertation is structured as follows. Chapter 2 provides a detailed exposition of spline spaces, spline estimators and their global asymptotic properties. Chapter 3 contains the known results on the local asymptotic properties of splines and is intended to motivate our further investigation of penalized splines and the choice of methods we use. Chapter 4 establishes the unified framework for all spline estimators and presents our results.

2 Spline estimators

Consider the nonparametric regression model for data pairs (x_i, Y_i)

$$Y_i = f(x_i) + \epsilon_i,\ i = 1, \ldots, N, \tag{2.1}$$

with standard assumptions on random errors

$$E(\epsilon_i) = 0,\ E(\epsilon_i \epsilon_j) = \sigma^2 \delta_{ij},\ \sigma^2 > 0, \tag{2.2}$$

where the δ_{ij} as the Kronecker delta, the design points $x_i \in [0, 1]$ are deterministic and f is an unknown, sufficiently smooth regression function. In this work, spline-based estimators of f will be considered.

In Section 2.1, different spline spaces and their bases are defined. In Section 2.2, we discuss asymptotic characteristics of nonparametric estimators. In Section 2.3 and 2.4, we give definitions and known global asymptotic properties of a smoothing and a regression spline estimator of f, respectively. In Section 2.5, we consider a penalized spline estimator as a trade-off between a smoothing and a regression spline.

2.1 Definitions of spline functions

Splines are smoothly connected piecewise polynomials. The points on which the polynomials are connected are called *knots*. Splines are defined by the degree of the piecewise polynomials, by the position of the knots and by conditions on smoothness at knots. In this section, we give formal definitions of different kinds of splines and spline spaces.

2.1.1 Spline spaces

We list and characterize some basic spline spaces. We use the following notations for $p, K \in \mathbb{N}$:

1. $\mathcal{P}_p \coloneqq \{p(x) : \ p(x) = \sum_{i=0}^{p} c_i x^i, \ c_0, \dots, c_p, x \in \mathbb{R}\}$ is a *space of polynomials* of degree p (order $p-1$).

2. $\underline{\tau}_K = \{0 = \tau_0 < \tau_1 \dots \tau_{K-1} < \tau_K = 1\}$ is a *partition* of the interval $[0,1]$. We refer to $\tau_i, \ i = 1, \dots, K-1$ as *inner knots.*

3. $\mathcal{M} = (m_1, \dots, m_{K-1})$ with $1 \leq m_i \leq p+1, \ i = 1, \dots, K-1$ is a *mulplicity vector.*

A polynomial spline is a piecewise polynomial, whose smoothness is controlled by the multiplicity vector $\mathcal{M}$. The next definition gives the details.

Definition 1 *The space of polynomial splines of degree p, of mulplicity $\mathcal{M} = (m_1, \dots, m_{K-1})$ and knots $\underline{\tau}_K = \{0 = \tau_0 < \tau_1 \dots \tau_{K-1} < \tau_K = 1\}$ is defined as*

$$\begin{aligned}\mathcal{S}(p; \mathcal{M}; \underline{\tau}_K) \coloneqq \{ & s : \exists \ s_1, \dots, s_K \in \mathcal{P}_p : s(x) = s_i(x) \ \textit{for} \ x \in [\tau_{i-1}, \tau_i], \ i = 1, \dots K \\ & \textit{and} \ s_i^{(j)}(\tau_i) = s_{i+1}^{(j)}(\tau_i), \ j = 0, \dots, p - m_i, \ i = 1, \dots K-1\}.\end{aligned}$$

In Schumaker (2007, p. 110), it is proved that $\mathcal{S}(p; \mathcal{M}; \underline{\tau}_K)$ is a linear space with dimension $p + 1 + \sum_{i=1}^{K-1} m_i$. The next definition gives the smoothest space of piecewise polynomials that corresponds to $\mathcal{M} = (1, \dots, 1)$.

Definition 2 *If $\mathcal{M} = (1, \dots, 1)$, $\underline{\tau}_K = \{0 = \tau_0 < \tau_1 \dots \tau_{K-1} < \tau_K = 1\}$ then space $\mathcal{S}(p; \mathcal{M}; \underline{\tau}_K)$ is called the space of polynomial splines of degree p with simple knots at points $\underline{\tau}_K$ and is denoted by $\mathcal{S}(p; \underline{\tau}_K)$, where*

$$\mathcal{S}(p; \underline{\tau}_K) \coloneqq \left\{ s \in C^{p-1}[0,1] : \ s(x) \in \mathcal{P}_p \ \textit{on each interval} \ [\tau_i, \tau_{i+1}], \ i = 0, \dots, K-1 \right\}.$$

The linear space $\mathcal{S}(p; \underline{\tau}_K)$ has dimension $p + K$.
Natural splines arise as the solution of the smoothing spline minimization problem, which will be discussed in Section 2.3.

Definition 3 *The space of natural splines of degree $2q-1$ with simple knots at $\underline{\tau}_K = \{0 = \tau_0 < \tau_1 \ldots \tau_{K-1} < \tau_K = 1\}$ is defined as*

$$\mathcal{NS}(2q-1; \underline{\tau}_K) \coloneqq \{s \in \mathcal{S}(2q-1; \underline{\tau}_K) : s \mid_{[0,\tau_1]}, s \mid_{[\tau_{K-1},1]} \in \mathcal{P}_{q-1}\},$$

where $s \mid_{[0,\tau_1]}$ and $s \mid_{[\tau_{K-1},1]}$ denote the restriction of function s to domain $[0,\tau_1]$ and $[\tau_{K-1}, 1]$ respectively.

Hence, a natural spline is $2q-2$ times continuously differentiable piecewise polynomial of degree $q-1$ on the intervals $[0,\tau_1]$ and $[\tau_{K-1}, 1]$ and of degree $2q-1$ on $[\tau_1, \tau_{K-1}]$. Because of the $2q$ constraints on intervals $[0,\tau_1]$ and $[\tau_{K-1}, 1]$, the dimension of natural spline space $\mathcal{NS}(2q-1; \underline{\tau}_K)$ is $K-1$.
Most of our results were obtained in the following space, which is comparatively simple because its functions are periodic and its knots are equidistant.

Definition 4 *The space of periodic splines of degree $p > 0$ based on equidistant knots $\underline{\tau}_K = \{i/K\}_{i=0}^K$ is defined as*

$$\mathcal{S}_{\texttt{per}}(p; \underline{\tau}_K) \coloneqq \left\{s \in \mathcal{S}\left(p; \underline{\tau}_K\right) : s^{(j)}(0) = s^{(j)}(1) \ for \ j = 0, \ldots, p-1\right\}.$$

Obviously, any periodic spline $s \in \mathcal{S}_{\texttt{per}}(p; \underline{\tau}_K)$ can be extended to the whole real line with $s(x+l) = s(x)$, $l \in \mathbb{Z}$. Because of the p periodicity constraints, the space of periodic splines $\mathcal{S}_{\texttt{per}}(p; \underline{\tau}_K)$ has dimension K.

2.1.2 B-spline bases

In this section, we define different types of B-spline functions that are used in the construction of bases of the spline spaces given in Section 2.1.1.
There are several ways to define a B-spline, such as divided difference (de Boor, 1978, p. 108) and recursion (Schumaker, 2007). Here we give an explicit definition, using truncated polynomials (based on Schoenberg, 1973).

Definition 5 *Given points $\tau_j, \ldots, \tau_{j+p+1} \in \mathbb{R}$, let $(x)_+^p : x \mapsto \{\max(0,x)\}^p$, $p > 0$ and $(x)_+^0 : x \mapsto I_{\{x \geq 0\}}$, where I is the indicator function. A general B-spline of degree p based*

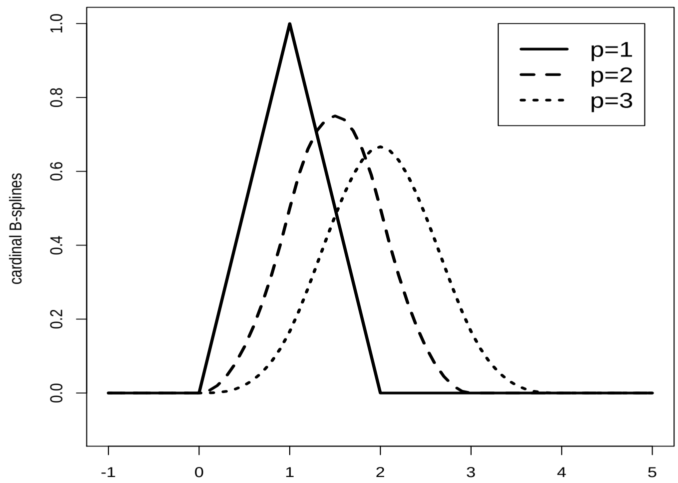

Figure 2.1: Forward cardinal B-spline $B_{c,p}(x)$ for $p = 1, 2, 3$

on $\tau_j, \dots \tau_{j+p+1}$ is defined as

$$\widetilde{B}_{j,p}(x) \coloneqq \sum_{i=0}^{p+1} \frac{(p+1)(\tau_{i+j} - x)_+^p}{\frac{\partial}{\partial x}\left\{(x-\tau_j)\dots(x-\tau_{j+p+1})\right\}_{x=\tau_{i+j}}}.$$

Functions $\widetilde{B}_{j,p}(x)$ have bounded support, that is $0 < \widetilde{B}_{j,p}(x)$ for $x \in (\tau_j, \tau_{j+p+1})$ and $\widetilde{B}_{j,p}(x) = 0$ for $x \notin (\tau_j, \tau_{j+p+1})$. Functions $\widetilde{B}_{j,p}(x)$ are normalized in that $\int_{-\infty}^{\infty} \widetilde{B}_{j,p}(x)dx = 1, \forall j, p$. Given a partition $\underline{\tau}_K = \{0 = \tau_0 < \tau_1 \dots \tau_{K-1} < \tau_K = 1\}$ of $[0,1]$, we define $2p$ additional knots $\tau_{-p} = \dots = \tau_{-1} = 0$ and $\tau_{K+1} = \dots = \tau_{K+p} = 1$. Then functions $\widetilde{B}_{j,p}(x),\ j = -p, \dots, K-1$ form a basis of $\mathcal{S}(p; \underline{\tau}_K)$. For more properties of a B-spline function, see de Boor (1978, chapter 9).

A particular case of the general B-splines are cardinal B-splines constructed on equidistant knots.

Definition 6 *A forward cardinal B-spline of degree p is defined as*

$$B_{c,p}(x) \coloneqq \frac{1}{p!} \sum_{i=0}^{p+1} (-1)^i \binom{p+1}{i} (x-i)_+^p, \tag{2.3}$$

where $(x)_+^p : x \mapsto \{\max(0,x)\}^p$ *, for* $p > 0$ *and* $(x)_+^0 : x \mapsto I_{\{x \geq 0\}}$*. A cardinal B-spline of degree* p *centered at* $i \in \mathbb{Z}$ *is defined as*

$$B_{c,p,i}(x) \coloneqq B_{c,p}(x + p/2 + 1/2 - i). \tag{2.4}$$

Taking in Definition 5 $\{\tau_j = j\}_{j=0}^{p+1}$, one can check that $B_{c,p}(x) = \widetilde{B}_{0,p}(x)$, that is $B_{c,p}(x)$ is a general B-spline of degree p based on knots $0, \dots, p+1$. The plot of $B_{c,p}(x)$ for different p is given in Figure (2.1). For more details about forward cardinal splines, we refer to Schoenberg (1973, lecture 3). Here, we mention only one well-known property we will use:

$$\int_{-\infty}^{\infty} B_{c,p}\left(x + \frac{p+1}{2}\right) \exp(-2\pi \mathtt{i} x) dx = \operatorname{sinc}(\pi x)^{p+1}, \tag{2.5}$$

where function $\operatorname{sinc}(x)$ is defined via

$$\operatorname{sinc}(x) \coloneqq \begin{cases} \sin(x)/x, & x \neq 0 \\ 1, & x = 0. \end{cases}$$

Finally, we define periodic B-splines on equidistant knots.

Definition 7 *The j-th periodic B-spline for the partition* $\underline{\tau}_K = \{i/K\}_{i=0}^K$ *is given by*

$$B_{j,p}(x) \coloneqq \sum_{l=-\infty}^{\infty} B_{c,p}\left\{K\left(x + l - j/K\right)\right\}. \tag{2.6}$$

Functions $B_{j,p}(x),\ j = 1, \dots, K$ build a basis of space $\mathcal{S}_{\mathtt{per}}(p; \underline{\tau}_K)$. Function $B_{j,p}(x)$ has period 1 and is a periodic extension to the real line of $\widetilde{B}_{j,p}(x)$, based on knots $\{\tau_i = i/K\}_{i=j}^{j+p+1}$. From (2.5), the Fourier series of periodic splines are given by

$$B_{i,p}(x) = \sum_{l=-\infty}^{\infty} \operatorname{sinc}(\pi l/K)^{p+1} \exp\{2\pi \mathtt{i} l(x - i/K)\},\ i = 1, \dots, K. \tag{2.7}$$

For $p = 3$ and $K = 4$, the periodic B-spline basis is plotted in Figure 2.2.

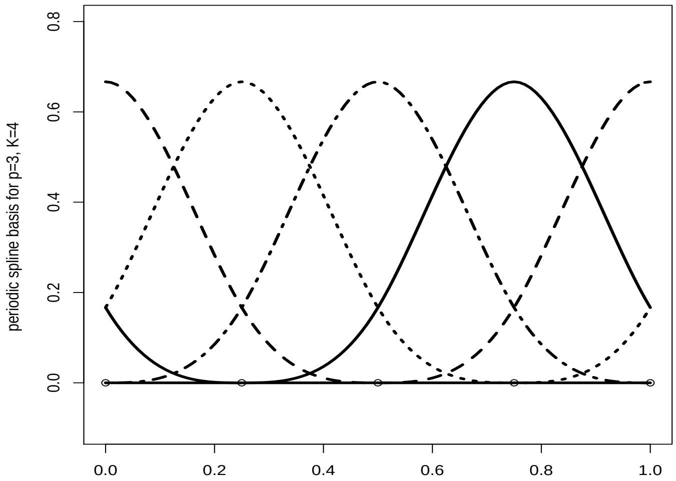

Figure 2.2: Periodic B-spline $B_{j,p}(x)$ for $p = 3$, $K = 4$

2.2 Global asymptotic characteristics of nonparametric estimators

In this section, we define the best achievable (optimal) rate of convergence of estimators and give it for different spaces. Further, we define measures of goodness of an estimator that allow us to check whether the estimator has the best possible rate of convergence or not.

Definition 8 *A positive sequence $\{a_N\}_{N=1}^{\infty}$ is called an optimal rate of convergence of estimators on class of functions $\mathcal{F}$ with respect to norm $\|\cdot\|$ if $\exists$ $c_1, c_2 > 0$*

$$\liminf_{N\to\infty} \inf_{\widehat{f}_N \in \widehat{\mathcal{F}}_{N,f}} \sup_{f\in\mathcal{F}} E\left[\left\|\widehat{f}_N - f\right\|^2 a_N^{-2}\right] \geq c_1$$
$$\limsup_{N\to\infty} \inf_{\widehat{f}_N \in \widehat{\mathcal{F}}_{N,f}} \sup_{f\in\mathcal{F}} E\left[\left\|\widehat{f}_N - f\right\|^2 a_N^{-2}\right] \leq c_2,$$

where $\widehat{\mathcal{F}}_{N,f}$ denotes a class of estimators of f based on random sample of size N.

Subsequently, we take $\widehat{\mathcal{F}}_{N,f}$ to be the set of all estimators that are linear in observations. The next lemma gives optimal rates of convergence for linear estimators with the deterministic design in Sobolev space.

Lemma 1 *(Speckman, 1985)*
In Definitionin 8, let $\mathcal{F} = W^q[0,1]$, *where*

$$\begin{aligned} W^q[0,1] &\coloneqq \Big\{ f : [0,1] \to \mathbb{R},\ f, f', \dots, f^{(q-1)} \textit{are absolutely continuous,} \\ &\quad \textit{and} \int_0^1 \left\{ f^{(q)}(x) \right\}^2 dx < \infty \Big\} \end{aligned}$$

is a Sobolev space of order q. For fixed $\{x_i\}_{i=1}^N, x_i \in [0,1]$ *with limiting density* $g(x)$ *such that* $\int_{x_i}^{x_{i+1}} g(x)dx = N^{-1}$, *let*

$$Y_{i,f} = f(x_i) + \epsilon_i,\ i = 1, \dots, N,\ f \in \mathcal{F},$$

where assumptions (2.1) holds for errors $\{\epsilon_i\}_{i=1}^N$. *Let* $\widehat{\mathcal{F}}_{N,f}$ *be all linear in* $\{Y_{i,f}\}_{i=1}^N$ *estimators. Under these assumptions, the optimal rate of convergence of estimators on* $\mathcal{F}$ *with respect to norm* $\|f\|_{N,2} = N^{-1} \left\{ \sum_{i=1}^N f(x_i)^2 \right\}^{-1/2}$ *is* $N^{-q/(2q+1)}$.

With additional assumptions on space $\mathcal{F}$ (for example, $\mathcal{F} = C^q[0,1]$ etc.) and error distribution, the results of Lemma 1 hold for the class of all estimators (including non-linear estimators) and random designs (see Tsybakov, 2009; Stone, 1980, 1982). Lemma 1 also holds in the Sobolev space for the class of all estimators (including non-linear) under additional assumptions on the error distribution (Golubev and Nussbaum, 1990).
The global goodness of estimator $\widehat{f}$ can be measured by the integrated squared mean error (the IMSE) that is defined for the equidistant design on interval $[0,1]$ via

$$\begin{aligned} IMSE(\widehat{f}) &\coloneqq E\left\|\widehat{f} - f\right\|_2^2 = E\int_0^1 \left\{\widehat{f}(x) - f(x)\right\}^2 dx \\ &= \int_0^1 \left\{E\widehat{f}(x) - f(x)\right\}^2 dx + \int_0^1 E\left\{\widehat{f}(x) - E\widehat{f}(x)\right\}^2 dx. \end{aligned} \tag{2.8}$$

The discrete version of the IMSE is the average mean squared error (the AMSE)

$$\begin{aligned} AMSE(\widehat{f}) &\coloneqq \frac{1}{N}\sum_{i=1}^{N}\left\{\widehat{f}(x_i)-f(x_i)\right\}^2 \qquad (2.9)\\ &= \frac{1}{N}\sum_{i=1}^{N}\left\{E\widehat{f}(x_i)-f(x_i)\right\}^2+\frac{1}{N}\sum_{i=1}^{N}E\left\{\widehat{f}(x_i)-E\widehat{f}(x_i)\right\}^2. \end{aligned}$$

Clearly, IMSE($\widehat{f}$) and AMSE($\widehat{f}$) are asymptotically equivalent for $N\to\infty$ under certain assumptions on the regularity of the data points. Estimator $\widehat{f}$ is said to be asymptotically optimal on $\mathcal{F}$ with respect to norm $\|\cdot\|_2$ if its IMSE (the AMSE) has the same order of magnitude as the corresponding squared optimal rate of convergence of estimators on $\mathcal{F}$. From Lemma 1, if the regression function is from $W^q[0,1]$, then asymptotically optimal linear estimators have the IMSE (the AMSE) of order $N^{2q/(2q+1)}$.

2.3 Smoothing spline estimators

In this section, we discuss smoothing splines and their global asymptotic properties. Smoothing splines are the oldest spline-based method. The idea of smoothing splines traces back to Whittaker (1923) and has been developed further by, among many others, Schoenberg (1964), Reinsch (1967) and Wahba (1975), who popularized smoothing splines in statistics. To obtain L_2-error bounds for smoothing spline estimators, different methods have been employed. Among them are Fourier techniques (Rice and Rosenblatt, 1981, 1983; Cogburn and Davis, 1974), the reproducing kernel Hilbert spaces framework (Craven and Wahba, 1979) and the asymptotic correspondence of the smoothing spline minimization problem to a certain boundary value problem (Utreras, 1983). Let us proceed with the details.

In Sobolev space $W^q[0,1]$ defined in Section 2.2, we consider a minimization problem

$$\min_{g\in W^q[0,1]}\left[\frac{1}{N}\sum_{i=1}^{N}\{Y_i-g(x_i)\}^2+\lambda\int_0^1\left\{g^{(q)}(x)\right\}^2dx\right],\ \lambda>0,\ q>0, \qquad (2.10)$$

where observations $0<x_1<x_2\dots x_{N-1}<x_N<1$ and data $\{Y_i\}_{i=1}^N$ are from regression model (2.1) with conditions (2.2). The solution of (2.10) $\widehat{f}_{ss}$ is called a smoothing spline estimator. The trade-off between the fidelity to the data and the smoothness of $\widehat{f}_{ss}$ is controlled by *roughness penalty* $\lambda \int_0^1 \left\{g^{(q)}(x)\right\}^2 dx$ with penalty parameter λ. If $\lambda = 0$, $\widehat{f}_{ss}$ interpolates $\{Y_i\}_{i=1}^N$. If $\lambda \to \infty$, $\widehat{f}_{ss}$ is a polynomial of degree $q-1$.

It is a well-known result that $\widehat{f}_{ss}$ is a natural spline of degree $2q-1$ with inner knots at $x_i,\ i = 1, \dots, N$ (see, e.g., Eggermont and LaRiccia, 2009, Chapter 19.3). Hence (2.10) is equivalent to the minimization in $(N+2q)$-dimensional spline space $\mathcal{S}\left(2q-1; \underline{\tau}_{N-1}\right)$, where $\underline{\tau}_{N+1} = \{0 = x_0 < x_1 < x_2 \dots x_N < x_{N+1} = 1\}$. Defining $4q-2$ additional knots $x_{-2q+1} = \dots = x_{-1} = 0$ and $x_{N+2} = \dots x_{N+2q} = 1$, we can construct basis vector $B(x) = \left\{\widetilde{B}_{-2q+1,2q-1}(x), \dots \widetilde{B}_{N,2q-1}(x)\right\}$ with B-spline $\widetilde{B}_{j,2q-1}(x)$ based on $x_j, \dots, x_{j+2q}$ for $j = -2q+1, \dots, N$ (see Definition 5). In matrix form, the smoothing spline estimator can be written as

$$\widehat{f}_{ss}(x) = B(x)(B^T B + \lambda D_q)^{-1} B^T Y, \tag{2.11}$$

where $Y = (Y_1, \dots Y_N)^T$, $B = \left\{B(x_1)^T, \dots B(x_N)^T\right\}^T$ is a $N \times (N+2q)$ design matrix, $D_q = \int_0^1 B^{(q)}(x)^T B^{(q)}(x) dx$ is a $(N+2q) \times (N+2q)$ matrix.

Let us touch a few aspects of the global asymptotic properties of smoothing splines. The AMSE of smoothing splines in the periodic case was studied by Wahba (1975).

Lemma 2 *(Wahba, 1975)*
Suppose that $f \in W^{2q}_{\mathtt{per}}[0,1] = \left\{g \in W^{2q}[0,1], g^{(j)}(0) = g^{(j)}(1),\ j = 0, \dots, 2q-1\right\}$ *and its Fourier series are* $f(t) = \sum_{l=-\infty}^{\infty} \widetilde{f}_l \exp(2\pi \mathtt{i} lt)$. *Let* $\widehat{f}_{ss}^{\mathtt{per}}$ *be the solution of*

$$\min_{g \in W^{2q}_{\mathtt{per}}[0,1]} \left[\frac{1}{N}\sum_{i=1}^N \{Y_i - g(x_i)\}^2 + \lambda \int_0^1 \left\{g^{(q)}(x)\right\}^2 dx\right],$$

where $\{x_i = i/N\}_{i=1}^N$. *Then*

$$AMSE\left(\widehat{f}_{ss}^{\mathtt{per}}\right) = \left[\sum_{i=1}^{(N-1)/2} \frac{2\lambda^2 (2\pi i)^{4q} |\widetilde{f}_i|^2}{\left\{1 + \lambda(2\pi i)^{2q}\right\}^2} + \frac{\sigma^2 \int_0^\infty \left(1 + t^{2q}\right)^{-2} dt}{N\lambda^{1/2q}} + \frac{\sigma^2}{N}\right]\{1 + o(1)\}. \tag{2.12}$$

To obtain the results of Lemma 2, Wahba (1975) worked with the Fourier trigonometric basis. In the non-periodic case, another orthogonal basis is used. We give a brief exposition of it. As was mentioned above, the solution $\widehat{f}_{ss}$ is a natural spline in space

$\mathcal{NS}(2q-1;\underline{\tau}_{N+1})$, where $\underline{\tau}_{N+1} = \{0 = x_0 < x_1 < x_2 \ldots x_{N-1} < x_N < x_{N+1} = 1\}$. Demmler and Reinsch (1975) introduced a real-valued orthogonal basis $\{\varphi_{N,i}(x)\}_{i=1}^N$ of $\mathcal{NS}(2q-1;\underline{\tau}_{N+1})$ that is determined by conditions

$$\frac{1}{N}\sum_{l=1}^{N}\varphi_{N,i}(x_l)\varphi_{N,j}(x_l) = \delta_{i,j} \qquad \int_0^1 \varphi_{N,i}^{(q)}(x)\varphi_{N,j}^{(q)}(x)dx = \delta_{i,j}\nu_{N,i} \qquad (2.13)$$
$$\text{with} \qquad 0 = \nu_{N,1} = \ldots = \nu_{N,q} < \nu_{N,q+1} \leq \ldots \leq \nu_{N,N}$$

for $i,j = 1,\ldots,N$. In the Demmler-Reinsch basis, the solution of smoothing spline minimization problem (2.10) is

$$\widehat{f}_{ss}(x) = \frac{1}{N}\sum_{l,i=1}^{N}\frac{\varphi_{N,i}(x_l)\varphi_{N,i}(x)}{1+\lambda\nu_{N,i}}Y_l = \sum_{i=1}^{N}\frac{f_{N,i}\varphi_{N,i}(x)}{1+\lambda\nu_{N,i}} \qquad (2.14)$$
$$\text{with } f_{N,i} = \frac{1}{N}\sum_{l=1}^{N}\varphi_{N,i}(x_l)Y_l.$$

A continuous analogue of the Demmler-Reinsch basis $\{\varphi_i(x)\}_{i=1}^{\infty}$ is determined by

$$\int_0^1 \varphi_j(x)\varphi_i(x)dx = \delta_{i,j} \qquad \int_0^1 \varphi_i^{(q)}(x)\varphi_j^{(q)}(x)dx = \nu_j\delta_{i,j} \qquad (2.15)$$
$$\text{with} \qquad 0 = \nu_1 = \ldots = \nu_q < \nu_{q+1} \leq \nu_{q+2}\ldots$$

for $i,j \in \mathbb{N}$. From, e.g., Speckman (1985), it is known that functions $\{\varphi_i(x)\}_{i=1}^{\infty}$ build a complete orthogonal system in $W^q[0,1]$ under L_2-inner product. Hence, $\forall f \in W^q[0,1]$

$$f(x) = \sum_{l=1}^{\infty} f_l\varphi_l(x), \text{ with } f_l = \int_0^1 f(x)\varphi(x)dx \text{ and } \sum_{l=1}^{\infty} f_l^2\nu_l < \infty.$$

Using Demmler-Reinsch representation (2.14) of $\widehat{f}_{ss}$ and approximating $\lim_{N\to\infty}\varphi_{N,i}(x) = \varphi_i(x)$, $\lim_{N\to\infty}\nu_{N,i}(x) = \nu_i$ (see, e.g., Speckman, 1985), the IMSE of smoothing splines in the non-periodic case can be written as

$$IMSE(\widehat{f}_{ss}) = \left[\sum_{i=q+1}^{\infty}\frac{\lambda^2\nu_i^2 f_i^2}{(1+\lambda\nu_i)^2} + \frac{1}{N}\sum_{i=q+1}^{\infty}\frac{\sigma^2}{(1+\lambda\nu_i)^2} + \frac{q\sigma^2}{N}\right]\{1+o(1)\}. \qquad (2.16)$$

Let us consider the asymptotic rate of $IMSE(\widehat{f}_{ss})$. The second and the third summands in (2.16) are the integrated variance with the order of magnitude $N^{-1}\lambda^{-1/(2q)}$, where $\nu_i \asymp i^{2q}, i > q$ (Birkhoff, 1908) and an approximation by an integral was used (see, e.g., Speckman, 1985). Here and subsequently, the notation $a \asymp b$ denotes that $a = \text{const}\ b$, where const > 0 is some generic constant. The first sum in (2.16) is the squared integrated bias and its rate depends on the decay of f_i. The decay rate of f_i for $f \in W^{2q}[0,1]$ was studied by Utreras (1988), and he obtained for $i > l$ that

$$|f_i| \leq \left| \sum_{j=q}^{2q-1} O\left\{ \left|f^{(j)}(1)\right| + \left|f^{(j)}(0)\right| \right\} \nu_i^{-(j+1)/(2q)} + \nu_i^{-1} \int_0^1 \varphi_i(x) f^{(2q)}(x)dx) \right|. \tag{2.17}$$

From (2.17), the decay rate of f_i and, therefore, the asymptotic rate of the IMSE depend on the natural boundary conditions

$$f^{(j)}(0) = f^{(j)}(0) = 0,\ j = q, \ldots, 2q-1. \tag{2.18}$$

If (2.18) holds, the integrated squared bias can be bounded by

$$\int_0^1 \left| f(x) - E\widehat{f}_{ss}(x) \right|^2 dx \leq \lambda^2 \sum_{i=q+1}^{\infty} \nu_i^2 f_i^2 = \lambda^2 \int_0^1 \left\{ f^{(2q)}(x) \right\}^2 dx = O(\lambda^2).$$

Therefore, for $f \in W^{2q}[0,1]$ with natural boundary conditions (2.18), the global asymptotic error of smoothing splines is

$$IMSE(\widehat{f}_{ss}) = O(\lambda^2) + O\left\{ N^{-1}\lambda^{-1/(2q)} \right\} \tag{2.19}$$

with optimal $\lambda \asymp N^{-2q/(1+4q)}$ providing $IMSE(\widehat{f}_{ss}) = O\{N^{-4q/(4q+1)}\}$.
If, for example, the j-th natural boundary condition in (2.18) does not hold, the squared integrated bias becomes (Utreras, 1988)

$$\int_0^1 \left| f(x) - E\widehat{f}_{ss}(x) \right|^2 dx = O\left\{ \lambda^{(2j+1)/(2q)} \right\},$$

where (2.17), $\nu_i \asymp i^{2q}, i > q$ and an approximation of the first term in (2.16) by an integral were used. Therefore, the IMSE has a bigger order of magnitude than in (2.19) and for $f \in W^{2q}[0,1]$ with $f^{(j)}(0)^2 + f^{(j)}(0)^2 \neq 0$ for some $j \in \{q, \ldots, 2q-1\}$ the global asymptotic error of smoothing splines is

$$IMSE(\widehat{f}_{ss}) = O\{\lambda^{(2j+1)/(2q)}\} + O\left\{N^{-1}\lambda^{-1/(2q)}\right\} \tag{2.20}$$

with optimal $\lambda \asymp N^{-q/(j+1)}$ providing $IMSE(\widehat{f}_{ss}) = N^{-(2j+1)/(2j+2)}$.
If f does not satisfy the natural boundary conditions, the IMSE is dominated by contribution from the boundary. If the IMSE (the AMSE) of an estimator is dominated by the contributions from the boundary, we say that the estimator has *boundary effects*.
For $f \in W^q[0,1]$, the global asymptotic properties were studied in Craven and Wahba (1979); Cox (1983); Eggermont and LaRiccia (2009). The IMSE of smoothing splines for $f \in W^q[0,1]$ is

$$IMSE\left(\widehat{f}_{ss}\right) = O(\lambda) + O\left\{N^{-1}\lambda^{-1/(2q)}\right\}.$$

with optimal $\lambda \asymp N^{-2q/(1+2q)}$ providing $IMSE(\widehat{f}_{ss}) = N^{-2q/(1+2q)}$. Since in the interior the pointwise squared bias is of order $O(\lambda^2)$ (see Section 3.2.2), the boundary bias of $\widehat{f}_{ss}$ dominates the IMSE. Thus, smoothing splines have boundary effects.

2.4 Regression spline estimators

In this section, we study regression splines and their global asymptotic properties. Regression (or least-squares) splines are introduced in works of Hartley (1961) and Hudson (1966), among others. Global asymptotics of regression splines were studied in Agarwal and Studden (1980) and Zhou et al. (1998). Both make use of the results of Barrow and Smith (1978), who found a sharp estimate of the error for the best L_2-approximation of a smooth function by a splines set. Let us look more closely at the details.
We consider regression model (2.1) with error assumptions (2.2). The regression spline estimator $\widehat{f}_{rs}$ of degree p based on partition $\underline{\tau}_K = \{0 = \tau_0 < \tau_1 \ldots \tau_{K-1} < \tau_K = 1\}$ is defined to be the minimizer of

$$\min_{s \in S(p;\underline{\tau}_K)} \sum_{i=1}^{N} \{Y_i - s(x_i)\}^2. \tag{2.21}$$

The matrix form representation of the regression spline estimator is

$$\widehat{f}_{rs}(x) = B(x)(B^T B)^{-1} B^T Y, \tag{2.22}$$

where the $N \times (K+p)$ basis matrix $B = \left\{B(x_1)^T, \dots B(x_N)^T\right\}^T$ has colons $B(x) = \left\{\widetilde{B}_{-p,p}(x), \dots \widetilde{B}_{K-1,p}(x)\right\}$. For the construction of basis vector $B(x)$ see Section 2.1.2. The global asymptotics (the IMSE) of the regression spline estimator was studied in Agarwal and Studden (1980). For sake of simplicity, we present their result for equidistant knots and observations.

Lemma 3 *(Agarwal and Studden, 1980)*
Let $f \in C^{p+1}[0,1]$, $x_i = i/N,\ i = 1, \dots, N$, $\underline{\tau}_K = \{i/K,\ i = 0, \dots, K\}$. The solution $\widehat{f}_{rs}$ of (2.21) has the following IMSE

$$IMSE\left(\widehat{f}_{rs}\right) = \left[\frac{\mathcal{B}_{2p+2}}{(2p+2)!K^{2p+2}} \int_0^1 \left\{f^{(p+1)}(x)\right\}^2 dx + \sigma^2 K/N\right] \{1 + o(1)\},$$

where $\mathcal{B}_{2p+2}$ is the $(2p+2)$-th Bernoulli number.

The first term of the IMSE is the integrated squared bias, the second term is the integrated variance. The optimal rate of convergence $N^{-(p+1)/(2p+3)}$ is provided by $K \asymp N^{1/(2p+3)}$.

Unlike the smoothing spline estimator, the regression spline estimator does not have boundary effects. Zhou et al. (1998) showed that the same rate of convergence as in Lemma 3 is achieved locally for any $x \in [0,1]$, that is $\widehat{f}_{rs}(x) - f(x) = O_p\left\{N^{-(p+1)/(2p+3)}\right\}$ uniformly (see also Section 3.3.2 for details). However, the variance at the boundary of $[0,1]$ is much larger than in the interior, since there are fewer observations at the boundary (Zhou et al., 1998).

2.5 Penalized spline estimators

A trade-off between smoothing and regression spline estimators are penalized splines, which have become increasingly popular over last two decades, see Ruppert et al. (2003). The asymptotic properties of penalized spline estimators have got attention only recently. It has been discussed in Claeskens et al. (2009) that depending on the number of

knots taken, penalized splines have asymptotic behavior similar either to regression or to smoothing splines. Kauermann et al. (2009) studied in more detail the "small" number of knots scenario in the generalized regression context. Recently, Wang et al. (2011) have shown that in the asymptotic scenario with the "large" number of knots, the equivalent kernel for penalized splines is asymptotically equivalent to that of smoothing splines. All these works used mixed approaches, combining techniques for regression and spline estimators, depending on the asymptotic scenario.
This section introduces penalized splines as a smoothing technique and their global asymptotic properties.

Definition 9 *Given regression model (2.1) with error assumptions (2.2). Penalized spline estimator $\widehat{f}_{ps}$ of degree p based on partition $\underline{\tau}_K = \{0 = \tau_0 < \tau_1 \dots \tau_{K-1} < \tau_K = 1\}$ is the solution of the following minimization problem:*

$$\min_{s \in \mathcal{S}(p;\underline{\tau}_K)} \left[\frac{1}{N} \sum_{i=1}^{N} \{Y_i - s(x_i)\}^2 + \lambda \int_0^1 \left\{ s^{(q)}(x) \right\}^2 dx \right], \ \lambda > 0, \ 0 < q \leq p. \tag{2.23}$$

If $\lambda = 0$, penalized spline estimator $\widehat{f}_{ps}$ coincides with regression spline estimator $\widehat{f}_{rs}$, the solution of (2.21). The penalized spline estimator of degree $p = 2q - 1$ based on partition $\underline{\tau}_{N+1} = \{0 = x_0 < x_1 < x_2 \dots < x_N < x_{N+1} = 1\}$ equals the smoothing spline estimator $\widehat{f}_{ss}$, the solution of (2.10).
The matrix representation of the penalized spline estimator is

$$\widehat{f}_{ps}(x) = B(x)(B^T B + \lambda D_q)^{-1} B^T Y, \tag{2.24}$$

where $B = \left\{ B(x_1)^T, \dots B(x_N)^T \right\}^T$ is $N \times (K + p)$ design matrix with vector-function $B(x) = \left\{ \widetilde{B}_{-p,p}(x), \dots \widetilde{B}_{K-1,p}(x) \right\}$ and $D_q = \int_0^1 B^{(q)}(x)^T B^{(q)}(x) dx$ is $(K + p) \times (K + p)$ matrix. For the construction of basis vector $B(x)$ see Section 2.1.2.
It was shown in Claeskens et al. (2009) that the asymptotic behavior of the penalized spline estimator depends on some parameter k_q, which is proportional to the maximum of the eigenvalues of $\lambda (B^T B)^{-1} D_q$ and depends on λ, K, q.

Lemma 4 *(Claeskens et al., 2009)*
Let $k_q \asymp \lambda^{1/(2q)} K$ be the maximum eigenvalue of $\lambda (B^T B)^{-1} D_q$. The solution $\widehat{f}_{ps}$ of the

minimization problem (2.23) has the following global asymptotic properties:

$$AMSE(\widehat{f}_{ps}) = \begin{cases} O\left(\frac{K}{N}\right) + O\left(\lambda^2 K^{2q}\right) + O\left(K^{-2p-2}\right), & k_q = O(1),\ f \in C^{p+1}[0,1] \\ O\left(K^{-2q}\right) + O\left(\lambda\right) + O\left\{\lambda^{-1/(2q)} N^{-1}\right\}, & k_q \to \infty,\ f \in W^q[0,1]. \end{cases}$$

If $k_q = O(1)$, the optimal convergence rate $N^{-(p+1)/(2p+3)}$ is achieved when $K \asymp N^{1/(2p+3)}$ and $\lambda = O(N^{-\gamma})$ with $\gamma \leq (p+1-q)/(2p+3)$.
If $k_q \to \infty$, the optimal convergence rate $N^{-q/(1+2q)}$ is achieved when $\lambda = O\left\{N^{-2q/(2q+1)}\right\}$, where $\lambda N^{2q} \to \infty$ and $K \asymp N^{\nu}$ with $\nu \geq 1/(2q+1)$.

The AMSE consists of three summands that give the orders of magnitude of the average squared approximation bias (bias that arises due to the approximation of the regression function by splines), the average squared shrinkage bias (bias that arises due to the penalization) and the average variance, respectively.
In case $k_q = O(1)$, the asymptotically optimal number of knots K (its rate of magnitude) and the optimal convergence rate of $\widehat{f}_{ps}$ coincide with those of regression splines (see Section 2.5). For $\lambda \asymp N^{-(p+1+q)/(2p+3)}$ and for asymptotically optimal K, the average squared shrinkage bias contributes to the AMSE, while for λ of a smaller order of magnitude, the average squared shrinkage bias is asymptotically negligible (Claeskens et al., 2009).
In case $k_q \to \infty$, the asymptotically optimal order of magnitude of λ and the optimal convergence rate of $\widehat{f}_{ps}$ coincide with those of smoothing splines (see Section 2.3). For $K \asymp N^{1/(2q+1)}$ and for asymptotically optimal λ, the average squared approximation bias contributes to the AMSE, while for K of a bigger order of magnitude, the average squared approximation bias is asymptotically negligible (Claeskens et al., 2009).
Note that the optimal convergence rate of $\widehat{f}_{ps}$ is determined by p for $k_q = O(1)$ and by q for $k_q \to \infty$. Since we assume $q \leq p$, the optimal convergence rate of $\widehat{f}_{ps}$ is faster for $k_q = O(1)$. This phenomenom can be explained by the shrinkage bias, which dominates when $k_q \to \infty$ and is a lot bigger at the boundary (Claeskens et al., 2009).

3 Local asymptotic properties and equivalent kernel for spline estimators

In this chapter, we bring together results on the pointwise behavior of three types of spline estimators that we introduced in the previous chapter.

In Section 3.1, we discuss one more method of nonparametric regression – kernel regression. As we will see later, methods of kernel regression can be used in order to find the pointwise bias and variance of spline estimators. In the remainder of Section 3.1, we introduce the concept of an equivalent kernel and give a short summary on kernels for each spline estimator. In Section 3.2, we discuss the local asymptotic properties of smoothing splines. In order to do it, we look more closely at the corresponding differential equations and their Green's functions. In Section 3.3, we discussion the local asymptotic properties and the equivalent kernel for regression splines. In Section 3.4, we have compiled a few known results concerning the pointwise asymptotic behavior of penalized spline estimators.

In this chapter and further, $\{x\}$ denotes the fractional part of x, $\lfloor x \rfloor = x - \{x\}$ is the largest integer not greater than x. Function $\mathcal{B}_p(x)$ denotes a Bernoulli polynomial of degree p

$$\mathcal{B}_p(x) \coloneqq \sum_{i=0}^{p} \frac{1}{i+1} \sum_{j=0}^{i} (-1)^j \binom{i}{j} (x+j)^p.$$

Notation $\mathcal{B}_p = \mathcal{B}_p(0)$ is used for denoting the Bernoulli numbers. $\delta(x)$ denotes the Dirac delta function.

For simplicity of notation, we give most results for equidistant knots and observations.

3.1 Definition of the equivalent kernel

Kernel regression is a standard technique of nonparametric regression to estimate unknown function f in regression model (2.1). Methods similar to kernel regression are used in order to find the local asymptotic properties of spline estimators. The concept of the equivalent kernel is used for it.

3.1.1 Kernel regression: basics

This section contains definitions of kernel estimators, their pointwise asymptotic behavior and a few aspects of boundary kernels.

Definition 10 *Function $\mathtt{K}(x,t) : \mathbb{R}^2 \to \mathbb{R}$ is called a kernel of order $p+1$ on $\mathbb{R}$ if and only if it satisfies the moment conditions*

1. $\int_{-\infty}^{\infty} \mathtt{K}(x,t)t^l dt = \begin{cases} 1, & l = 0, \\ x^l, & l = 1, \dots p \end{cases}$
2. $\int_{-\infty}^{\infty} \mathtt{K}(x,t)t^{p+1} dt \neq x^{p+1}$
3. $\int_{-\infty}^{\infty} \mathtt{K}^2(x,t) dt < \infty.$

Given regression model (2.1) and kernel of order $p+1$ $\mathtt{K}(x,t)$, let us define *kernel estimator* $\widehat{f}_{\mathtt{K}}$ for equidistant observations points $\{i/N\}_{i=1}^N$ as

$$\widehat{f}_{\mathtt{K}}(x) := \frac{1}{Nh} \sum_{i=1}^{N} \mathtt{K}\left(\frac{x}{h}, \frac{i}{Nh}\right) Y_i, \tag{3.1}$$

where positive parameter $h \to 0$ is called a *bandwidth* and plays a role similar to the penalty parameter for smoothing splines, i.e controls the trade-off between the fidelity to the data and the smoothness of the estimator.

Further, we discuss the pointwise bias and variance of the kernel estimator. For simplicity, we consider a translation-invariant kernel estimator with equidistant observations.

Definition 11 *Kernel $\mathtt{K}(x,t)$ is called translation-invariant if there exists $\mathtt{K}(t)$ such that $\mathtt{K}(x,t) = \mathtt{K}(x-t)$.*

Hence, the translation invariant kernel estimator can be written as

$$\widehat{f}_{\mathtt{K}}(x) = \frac{1}{Nh}\sum_{i=1}^{N} \mathtt{K}\left(\frac{x-i/N}{h}\right) Y_i \tag{3.2}$$

with the following moment conditions:

$$\int_{-\infty}^{\infty} \mathtt{K}(t)t^l dt = \begin{cases} 1, & l=0 \\ 0, & l=1,\dots p \end{cases}; \quad \int_{-\infty}^{\infty} \mathtt{K}(t)t^{p+1}dt \neq 0; \quad \int_{-\infty}^{\infty} \mathtt{K}^2(t)dt < \infty. \tag{3.3}$$

The local asymptotic properties of the translation-invariant kernel estimator are relatively easy to study. With some additional assumptions on the kernel, the technique for finding the pointwise bias and variance is similar in the non translation-invariant case. We refer to Section 4.4.4, wherein this case was considered. We now show how to find the pointwise bias and variance of the translation-invariant kernel estimator. The general idea for finding the pointwise bias is to use the Taylor series of $f \in C^{p+1}[0,1]$ for $t \in [0,1]$, $h \to 0$:

$$f(x-th) = f(x) - thf^{'}(x) + \ldots + \frac{(-th)^p f^{(p)}(x)}{p!} + \frac{(-th)^{p+1} f^{(p+1)}(x)}{(p+1)!} + o(h^{p+1}).$$

If $x \in (h, 1-h)$, interval $(x/h-h, x/h)$ is approximately equal to $(-\infty, \infty)$, since $h \to 0$. Therefore, conditions (3.3) on the moments can be applied:

$$\begin{aligned} E\widehat{f}_{\mathtt{K}}(x) &= \frac{1}{Nh}\sum_{i=1}^{N} \mathtt{K}\left(\frac{x-i/N}{h}\right) f(i/N) = h^{-1}\int_0^1 \mathtt{K}\left(\frac{x-t}{h}\right) f(t)dt + O(N^{-1}) \\ &= \int_{x/h-1/h}^{x/h} \mathtt{K}\left(t\right) f(x-th)dt + O(N^{-1}) \approx \int_{-\infty}^{\infty} \mathtt{K}\left(t\right) f(x-th)dt + O(N^{-1}) \\ &= f(x) + \frac{(-h)^{p+1} f^{(p+1)}(x)}{(p+1)!}\int_{-\infty}^{\infty} \mathtt{K}\left(t\right) t^{p+1}dt + o(h^{p+1}) + O(N^{-1}). \end{aligned} \tag{3.4}$$

The difference $\int_{x/h-1/h}^{x/h} \mathtt{K}\left(t\right) f(x-ht)dt - \int_{-\infty}^{\infty} \mathtt{K}\left(t\right) f(x-ht)dt$ depends on the particular kernel $\mathtt{K}(t)$, but in any case is asymptotically negligible for $h \to 0$.

The pointwise variance of the kernel estimator can be derived from

$$\begin{aligned}\operatorname{Var}\left\{\widehat{f}_{\mathtt{K}}(x)\right\} &= \frac{\sigma^2}{N^2h^2}\sum_{i=1}^{N}\mathtt{K}^2\left(\frac{x-i/N}{h}\right) = \frac{\sigma^2}{Nh^2}\int_0^1 \mathtt{K}^2\left(\frac{x-t}{h}\right)dt + O(N^{-1}) \\ &= \frac{\sigma^2}{Nh}\int_{(x-1)/h}^{x/h}\mathtt{K}^2\left(t\right)dt + O(N^{-1}) = O(N^{-1}h^{-1}). \end{aligned} \tag{3.5}$$

With this, in the interior, the bias of the estimator $\widehat{f}_{\mathtt{K}}$ has order of magnitude h^{p+1}, while the variance has order $N^{-1}h^{-1}$. Thus, the bandwidth controls the trade-off between the bias and the variance: decreasing the pointwise bias, we increase the pointwise variance. Kernel regression estimator $\widehat{f}_{\mathtt{K}}$ has a bias of a bigger order at the boundary than in the interior, where the boundary region is $[0,h)$ and $(1-h,1]$. There are many ways to correct the boundary bias, such as the boundary correction kernel method. The idea is to use a corrected kernel that satisfies certain incomplete moment conditions at the boundary, so that the same technique as in (3.4) can be applied for the boundary points. For $x=\alpha h,\alpha \in [0,1)$, these conditions on boundary kernel $\mathtt{K}_\alpha(t)$ are

$$\int_{-\alpha}^{\infty}\mathtt{K}_\alpha(t)t^l dt = \begin{cases}1, & l=0\\ 0, & l=1,\dots p\end{cases}; \qquad \int_{-\alpha}^{\infty}\mathtt{K}_\alpha(t)t^{p+1}dt \neq 0.$$

For $x=1-\alpha h,\alpha \in [0,1)$, the conditions on boundary kernel $\mathtt{K}_\alpha(t)$ are

$$\int_{-\infty}^{\alpha}\mathtt{K}_\alpha(t)t^l dt = \begin{cases}1, & l=0\\ 0, & l=1,\dots p\end{cases}; \qquad \int_{-\infty}^{\alpha}\mathtt{K}_\alpha(t)t^{p+1}dt \neq 0.$$

Kernel $\mathtt{K}_\alpha(t)$ can be constructed for each $x\in[0,h)$ or $x\in(1-h,1]$ by weighting the original kernel $\mathtt{K}(t)$ with polynomials so that the resulting function satisfies the incomplete moment conditions above. The coefficients of the polynomials can be expressed in terms of incomplete moments of function $\mathtt{K}(t)$ (see Gasser and Müller, 1979; Gasser and Müller, 1984). In the literature (e.g., Gasser and Müller, 1979), kernel $\mathtt{K}(t)$ and $\mathtt{K}_\alpha(t)$ are usually assumed to have compact support and the integrals in the moment conditions are typically taken over a finite interval.

Finally, we discuss one more aspect of the global asymptotic properties for kernel estimator $\widehat{f}_{\mathtt{K}}$ from (3.1). In general, the moment conditions in Definition 10 are not sufficient in order to provide the optimal rate of convergence for estimator $\widehat{f}_{\mathtt{K}}$. The following lemma gives sufficient conditions on kernel $\mathtt{K}(x,t)$ under which kernel regression estimator $\widehat{f}_{\mathtt{K}}$ achieves the optimal convergence rate.

Lemma 5 *Eggermont and LaRiccia (2009, theorem 14.2.16)*
For $f \in W^q[0,1]$ and $\widehat{f}_{\mathtt{K}}$ from (3.1), $IMSE(\widehat{f}_{\mathtt{K}}) = O\left\{N^{-2q/(2q+1)}\right\}$ if $h \asymp N^{-1/(2q+1)}$ and $h^{-1}\mathtt{K}(h^{-1}x, h^{-1}t)$ is convolution-like kernel of order q, that is

1. $\forall x \in [0, h^{-1}]\ \exists c > 0$

$$\int_0^{h^{-1}} |\mathtt{K}(x,t)|\, dt \le c; \qquad \sup_{t\in[0,h^{-1}]} |\mathtt{K}(x,t)| \le c; \qquad \int_0^{h^{-1}} \left|\frac{\partial}{\partial t} K(x,t)\right| dt \le c.$$

2. $\forall x \in [0, h^{-1}]\ \exists c > 0$

$$\int_0^{h^{-1}} (x-t)^l \mathtt{K}\,(x,t)\, dt = \begin{cases} 1, & l = 0, \\ 0 & l = 1, \dots, q-1 \end{cases}$$
$$\int_0^{h^{-1}} |x-t|^q \left|\mathtt{K}\,(x,t)\right| dt \le c.$$

For more global asymptotic properties of the kernel estimator we refer to Eggermont and LaRiccia (2009).

3.1.2 Equivalent kernel for spline estimators: definitions

The idea of an equivalent kernel for spline estimators is to provide a simpler estimator than the original one. In this section, we discuss an equivalent kernel for all three spline estimators together and an equivalent kernel for each of them in particular.
From matrix-form representations (2.11), (2.22) and (2.24), it follows that spline estimators are linear in observations $Y_1, \dots, Y_N$, i.e., there is the unique weight function $W(x,t)$ such that spline estimator $\widehat{f}$ can be represented in the form

$$\widehat{f}(x) = \frac{1}{N}\sum_{i=1}^{N} W(x, x_i) Y_i. \tag{3.6}$$

In general, function $W(x,t)$ depends on the design points $x_1, \dots, x_N$, on the corresponding spline space (its degree, the number of knots and the position of the knots), on the penalty parameters (in case of smoothing and penalized splines). Function $W(x,t)$ is called the *effective kernel*. We consider an equivalent estimator

$$\widetilde{f}(x) = \frac{1}{N} \sum_{i=1}^{N} \widetilde{W}(x, x_i) Y_i$$

such that function $\widetilde{W}(x,t)$ does not depend on the design points and $\forall x \in [0,1]$

$$\left| \widehat{f}(x) - \widetilde{f}(x) \right| = o_p \left\{ E\widehat{f}(x) - f(x) \right\}.$$

We call function $\widetilde{W}(x,t)$ an *equivalent kernel*. We also refer to scaled versions of $\widetilde{W}(x,t)$ as equivalent kernels. Equivalent estimator $\widetilde{f}$ is simpler than original spline estimator $\widehat{f}$ and, in many cases, allows to study the pointwise asymptotic behavior of spline estimators using the methods of kernel regression discussed in Section 3.1.1.

In the following, we give a literature overview of equivalent kernels for each spline estimator.

The equivalent kernel for smoothing splines is the Green's function of the corresponding continuous differential equation. For the discussion see, e.g., Eggermont and LaRiccia (2009) and Section 3.2.1. This Green's function can be approximated by a translation-invariant function in the interior that has a simple explicit form. At the boundary, the form of the equivalent kernel (Green's function) is more complicated, but its behavior is studied and the explicit form can be obtained in many cases.

The equivalent kernel for regression splines is an L_2-projection kernel on a certain linear spline space (see Huang and Studden, 1992). The equivalent kernel is not translation-invariant and its shape depends on the distance from the knots. Its formula is obtained in terms of B-splines and roots of some polynomials, which can be computed at least for $p = 1, 2, 3$. For more details see Huang and Studden (1992) and Section 3.3.1.

The equivalent kernels for penalized splines is usually associated with the equivalent kernel for smoothing splines. That is, the known results are obtained only for the $k_q \to \infty$ scenario (see Section 2.5 for the asymptotic scenarios and Section 3.4 for an overview of the literature). We treat different asymptotic scenarios and obtain new results that establish the dependence of the equivalent kernel on k_q in Section 4.3.

3.2 Equivalent kernel for smoothing splines

To understand the local properties of smoothing spline estimators, asymptotic equivalent kernels have been studied extensively. First, Cogburn and Davis (1974) obtained an asymptotic equivalent kernel for smoothing splines on the real line, using Fourier techniques. Messer and Goldstein (1993) and Thomas-Agnan (1996) extended this kernel to the case of a bounded interval. Later, Eggermont and LaRiccia (2006) refined these two results. In this section, we look more closely at the equivalent kernel for smoothing splines. Then, we define with its help equivalent estimator $\widetilde{f}_{ss}$ and study the local asymptotic properties of $\widetilde{f}_{ss}$.

3.2.1 Formula of the equivalent kernel for smoothing splines

The problem of finding an approximation of the weight function in (3.6) for smoothing spline estimator $\widehat{f}_{ss}$ given in Section 2.3 is known and has been treated exhaustively to the verge of being solved. The general idea is to approximate $\widehat{f}_{ss}$ by the solution of some continuous minimization problem, which can be solved using differential equations and Fourier analysis. In Section 3.2.1, we first discuss smoothing spline related minimization problems and their corresponding boundary value problems, then we define the equivalent kernel and examine its properties. Finally, we study the behavior of the equivalent kernel in the interior.

From Eggermont and LaRiccia (2009, p.63), the Euler equation of smoothing spline minimization problem (2.10) is

$$\begin{aligned} \frac{1}{N}\sum_{i=1}^{N} g(x_i)\delta(x-x_i) + \lambda(-1)^q g^{(2q)}(x) &= \tfrac{1}{N}\textstyle\sum_{i=1}^{N} Y_i\delta(x-x_i) \\ g^{(j)}(0) = g^{(j)}(1) = 0, &\qquad j = q, \dots, 2q-1. \end{aligned} \tag{3.7}$$

The standard approach is to approximate the discrete sums in (3.7) by integrals and to look for the solution of the following Euler equation

$$\begin{aligned} g(x) + \lambda(-1)^q g^{(2q)}(x) &= f(x),\ x \in (0,1) \\ g^{(j)}(0) = g^{(j)}(1) &= 0,\ j = q, \dots 2q-1 \end{aligned} \tag{3.8}$$

that corresponds to Tikhonov regularization problem

$$\min_{g \in W^q[0,1]} \left[\int_0^1 \{g(x) - f(x)\}^2 \, dx + \lambda \int_0^1 \left\{ g^{(q)}(x) \right\}^2 dx \right]. \tag{3.9}$$

A slightly different approach, to be found in Eggermont and LaRiccia (2009), is to approximate only the first discrete sum in (3.7) and work with solution $\widetilde{f}$ of the following Euler equation

$$\begin{aligned} g(x) + \lambda(-1)^q g^{(2q)}(x) &= N^{-1} \textstyle\sum_{i=1}^N Y_i \delta(x - x_i) \\ g^{(j)}(0) = g^{(j)}(1) = 0, &\qquad j = q, \dots 2q-1, \end{aligned} \tag{3.10}$$

which corresponds to minimization problem

$$\min_{g \in W^q(0,1)} \left[\int_0^1 g^2(x) dx - \frac{2}{N} \sum_{i=1}^N Y_i g(x_i) + \lambda \int_0^1 \left\{ g^{(q)}(x) \right\}^2 dx \right]. \tag{3.11}$$

The last minimization problem corresponds to smoothing spline minimization problem (2.10) with only $N^{-1} \sum_{i=1}^N g(x_i)$ approximated by $\int_0^1 g^2(x) dx$.
To define the equivalent kernel, we need to introduce the concept of the Green's function and the reproducing kernel. The Green's function $\mathcal{R}_\lambda(x,t)$ for (3.8) and (3.10) satisfies

$$\mathcal{R}_\lambda(x,t) + \lambda(-1)^q \frac{\partial^{2q} \mathcal{R}_\lambda(x,t)}{\partial t^{2q}} = \delta(x - t), \tag{3.12}$$

$$\left[\frac{\partial^j \mathcal{R}_\lambda(x,t)}{\partial t^j} \right]_{t=0} = \left[\frac{\partial^j \mathcal{R}_\lambda(x,t)}{\partial t^j} \right]_{t=1} = 0 \quad ,\ j = q, \dots, 2q-1, x \neq 0, 1. \tag{3.13}$$

Now we give the definition of the reproducing kernel.

Definition 12 *Let $\mathcal{H}$ be a Hilbert space on $[0,1]$ with inner product $\langle f, g \rangle, f, g \in \mathcal{H}$.*

Function $\mathcal{R} : \mathcal{H}\times\mathcal{H}\to \mathbb{R}$ is called the reproducing kernel iff

1. $\mathcal{R}(x,t) \in \mathcal{H}$ *for fixed* $t \in [0,1]$
2. $g(x) = \langle \mathcal{R}(x,t), g(t)\rangle\,,\ \forall g \in \mathcal{H},\ \forall x \in [0,1]$.

The details on the Green's function can be found in, e.g., Stakgold (1979), while the details on the reproducing kernels can be found in (Berlinet and Thomas-Agnan, 2004). In Eggermont and LaRiccia (2009), it was shown that the Green's function of (3.8), (3.10) exists, is unique and equals the reproducing kernel of $W^q[0,1]$ with inner product $(f,g)_\lambda = \int_0^1 f(t)g(t)dt + \lambda\int_0^1 f^{(q)}(t)g^{(q)}(t)dt$. In the literature, function $\mathcal{R}_\lambda(x,t)$ is called the *equivalent kernel for smoothing spline* estimator $\widehat{f}_{ss}$.

By Green's theorem, the solutions of boundary value problems (3.8) and (3.10) are $\int_0^1 \mathcal{R}_\lambda(x,t)f(t)dt$ and $\int_0^1 \mathcal{R}_\lambda(x,t)N^{-1}\sum_{i=1}^N Y_i\delta(x-x_i)dt$, correspondingly. The last solution is denoted $\widetilde{f}_{ss}$ and was treated in detail in Eggermont and LaRiccia (2009). It can be written as

$$\begin{aligned}\widetilde{f}_{ss}(x) &= N^{-1}\sum_{i=1}^N \mathcal{R}_\lambda(x,i/N)Y_i \qquad (3.14)\\ &= \int_0^1 \mathcal{R}_\lambda(x,t)f(t)dt + N^{-1}\sum_{i=1}^N \mathcal{R}_\lambda(x,i/N)\epsilon_i + O(N^{-1}).\end{aligned}$$

The following bound from Eggermont and LaRiccia (2009) shows how good $\widetilde{f}$ approximates $\widehat{f}$

$$E\left\{\sup_{x\in[0,1]}\left|\widehat{f}_{ss}(x)-\widetilde{f}_{ss}(x)\right|^2\right\} = O\left\{\lambda^{(4q-1)/(2q)} + \lambda^{-3/(2q)}N^{-2}\right\}. \qquad (3.15)$$

Further, we find the explicit form for Green's function (equivalent kernel) $\mathcal{R}_{ss}(x,t)$. Finding the Green's function of system (3.8) and (3.10) is equivalent to finding the general solution of (3.12) as a linear combination of homogeneous solutions and matching coefficients so that they satisfy boundary conditions (3.13).

For $q = 1$, the Green's function was found in Nychka (1995)

$$\mathcal{R}_\lambda(x,t) = \frac{\exp\left(\lambda^{-1/2}|x-t|\right) + \exp\{\lambda^{-1/2}(-t-x)\}}{\lambda^{-1/2}\{2-2\exp(-\lambda^{-1/2})\}} + \frac{\exp\{\lambda^{-1/2}(t+x)-2)\} + \exp(2-\lambda^{-1/2}|t-x|)}{\lambda^{1/2}\{2-2\exp(-\lambda^{1/2})\}}.$$

For $q = 2$, the Green's function is already quite complicated (see, e.g., Thomas-Agnan, 1996). Hence, we give only the following form for all $q \geq 1$

$$\mathcal{R}_\lambda(x,t) = \lambda^{-1/(2q)}\mathcal{K}_{ss}\left\{\lambda^{-1/(2q)}(x-t)\right\} + \lambda^{-1/(2q)}\sum_{i=0}^{2q-1}\alpha_i\{\lambda^{-1/(2q)}t\}\psi_i\left\{\lambda^{-1/(2q)}x\right\}, \quad (3.16)$$

where $\lambda^{-1/(2q)}\mathcal{K}_{ss}\left\{\lambda^{-1/(2q)}(x-t)\right\}$ is the solution of (3.12) with boundary conditions (3.13) moved to infinity, $\psi_i\left\{\lambda^{-1/(2q)}x\right\}, i = 1,\ldots,2q-1$ are the homogeneous solutions of (3.12) and coefficients $\alpha_i\{\lambda^{-1/(2q)}t\}\, i = 1,\ldots,2q-1$ are chosen to match natural boundary conditions (3.13) for each t. Representation (3.16) is equivalent to $\mathcal{R}_\lambda(x,t) = \lambda^{-1/(2q)}\mathcal{K}_{ss}^{[0,1]}\{\lambda^{-1/(2q)}x, \lambda^{-1/(2q)}t\}$, where

$$\mathcal{K}_{ss}^{[0,1]}(x,t) \coloneqq \mathcal{R}_{\lambda=1}(x,t) = \mathcal{K}_{ss}(x-t) + \sum_{i=0}^{2q-1}\alpha_i(t)\psi_i(x) \quad (3.17)$$

is referred to as the (scaled) equivalent kernel and $\lambda^{1/(2q)}$ is the smoothing spline bandwidth. Let us discuss each term in (3.16) and (3.17) in more detail.
The homogeneous solutions of (3.12), $\psi_i\left\{\lambda^{-1/(2q)}x\right\}, i = 1,\ldots,2q-1$, are found via the roots of characteristic polynomial $t + \lambda(-1)^q t^{2q} = 0$. It is convenient to write them in the form

$$\psi_i(x) = \begin{cases} \exp\left\{\exp\left(\frac{2i+q+1}{2q}\pi\mathtt{i}\right)x\right\}, & i = 0,\ldots,q-1 \\ \exp\left\{\exp\left(\frac{2i+q+1}{2q}\pi\mathtt{i}\right)(x-1)\right\}, & i = q,\ldots 2q-1. \end{cases} \quad (3.18)$$

Coefficients $\alpha_i\{\lambda^{-1/(2q)}t\},\ i=0,\dots 2q-1$ decay exponentially away from the boundary. More precisely, from Eggermont and LaRiccia (2006), it holds that

$$\sup_{0\le i\le q-1}|\alpha_i(t)| = O\left\{\exp(-\alpha t)\right\},\quad \sup_{q\le i\le 2q-1}|\alpha_i(t)| = O\left\{\exp(-\alpha(1-t))\right\},\ \alpha\ >0. \tag{3.19}$$

Finally, let us study function $\mathcal{K}_{ss}(x)$ that plays an important role in our further analysis. From (3.16), $\lambda^{-1/(2q)}\mathcal{K}_{ss}\left\{\lambda^{-1/(2q)}(x-t)\right\}$ corresponds to the Green's function of Euler equation

$$\begin{aligned} g(x)+\lambda(-1)^q g^{(2q)}(x) &= f(x),\ x\in\mathbb{R} \\ \lim_{|x|\to\infty} g^{(j)}(x) &= 0,\ j=q,\dots 2q-1, \end{aligned} \tag{3.20}$$

which in turn corresponds to the minimization problem on the real line

$$\min_{g\in W^q(\mathbb{R})}\left[\int_{-\infty}^{\infty}\left\{g(x)-f(x)\right\}^2 dx+\lambda\int_{-\infty}^{\infty}\left\{g^{(q)}(x)\right\}^2 dx\right]. \tag{3.21}$$

From (3.21), the explicit form of $\mathcal{K}_{ss}(x)$ can be derived by means of Fourier transform (see Thomas-Agnan, 1996) and is given in Definition 13.

Definition 13 *Function $\mathcal{K}_{ss}:\mathbb{R}\to\mathbb{R}$ defined via integral*

$$\mathcal{K}_{ss}(x) \coloneqq \int_{-\infty}^{\infty}\frac{\exp(2\pi u x)}{1+(2\pi u)^{2q}}du \tag{3.22}$$

is called the equivalent kernel for smoothing spline estimator (2.10) in the interior.

Definition 13 is motivated by the fact that from (3.19) in the interior of $[0,1]$ the equivalent kernels can be approximated by $\mathcal{R}_\lambda(x,t)\approx\lambda^{-1/(2q)}\mathcal{K}_{ss}\left\{\lambda^{-1/(2q)}(x-t)\right\}$ and $\mathcal{K}_{ss}^{[0,1]}(x,t)\approx\mathcal{K}_{ss}(x-t)$. These approximations were first considered in Silverman (1985) for $q=2$. The integral (3.22) can be computed using the residue theorem (see Thomas-Agnan, 1996) and its closed-form expression is

$$\mathcal{K}_{ss}(x)=\sum_{j=0}^{q-1}\frac{\mathtt{i}\exp\left[\mathtt{i}\,|x|\exp\left\{\pi\mathtt{i}\,(2j+1)\,/(2q)\right\}\right]}{2q\exp\left\{\mathtt{i}\pi(2q-1)\,(2j+1)\,/(2q)\right\}}. \tag{3.23}$$

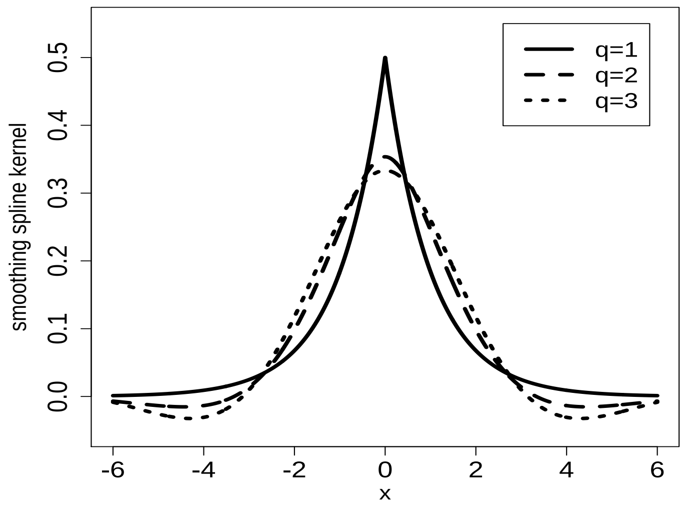

Figure 3.1: Equivalent kernel $\mathcal{K}_{ss}(x)$ for smoothing spline estimator

Formula (3.23) can be written without the complex exponentials by grouping the terms in the summation as

$$\mathcal{K}_{ss}(x) = \sum_{j=0}^{\lfloor q-1 \rfloor/2} \frac{\sin\left[\pi(2q-1)(2j+1)/(2q) - |x|\cos\{\pi(2j+1)/(2q)\}\right]}{q\exp\left[|x|\sin\{\pi(2j+1)/(2q)\}\right]} - \frac{I_{\{q\ is\ odd\}}}{2q\exp(|x|)}.$$

For fixed q, formula (3.23) becomes

$$\begin{aligned}
\mathcal{K}_{ss}(x) &= \frac{1}{2}\exp(-|x|),\ \ for\ q=1 \\
\mathcal{K}_{ss}(x) &= \frac{1}{2}\exp\left(-|x|/\sqrt{2}\right)\sin\left(|x|/\sqrt{2}+\frac{\pi}{4}\right),\ \ for\ q=2 \\
\mathcal{K}_{ss}(x) &= \frac{1}{6}\left\{\exp\left(-|x|\right)+2\exp\left(-|t|/2\right)\sin\left(|x|\frac{\sqrt{3}}{2}+\frac{\pi}{6}\right)\right\},\ \ for\ q=3.
\end{aligned}$$

The plot of $\mathcal{K}_{ss}(x)$ for $q = 1, 2, 3$ is given in Figure 3.1. As we can see, the equivalent kernel for smoothing splines in the interior of $[0, 1]$ is symmetric and decays very rapidly.
Remarks

1. The special periodic case was considered in Cogburn and Davis (1974). If $f \in$

$L_2[0,1]$ and $g \in W^q[0,1]$ are functions with period 1, then the solution of (3.21) has the form $\int_0^1 f(t)\mathtt{W}_{ss}(x-t)dt$ with

$$\mathtt{W}_{ss}(x) = \sum_{l=-\infty}^{\infty} \lambda^{-1/(2q)}\mathcal{K}_{ss}\left\{\lambda^{-1/(2q)}(x+l)\right\}.$$

2. Formulas 2.6 and 2.7 from Messer (1991) give an approximation of the Green's function for $q=2$ as a combination of $\mathcal{K}_{ss}(x)$ and its derivatives.

3. From Eggermont and LaRiccia (2009), Green's function $\mathcal{R}_\lambda(x,t)$ is a convolution-like kernel as defined in Lemma 5 with $h = \lambda^{1/(2q)}$.

3.2.2 Pointwise asymptotic behavior of smoothing splines

In this section, we study the pointwise bias and variance of smoothing spline estimator $\widehat{f}_{ss}$. From matrix form representation of $\widehat{f}_{ss}$, given in Section 2.3, the smoothing spline is linear in observations and can be written as

$$\widehat{f}_{ss}(x) = N^{-1}\sum_{i=1}^{N}\mathtt{W}_{ss}(x,i/N)Y_i,$$

where $\mathtt{W}_{ss}(x,t) = NB(x)\left(B^TB+\lambda D_q\right)^{-1}B^T(t)$ with matrices B, D_q and vector-basis $B(x)$ given in Section 2.3. From Nychka (1995), $\exists c_1, c_2 > 0$ such that

$$|\mathtt{W}_{ss}(x,t) - \mathcal{R}_{ss}(x,t)| < O\left[\frac{\lambda^{-1/(2q)}N^{-1}}{\lambda^{1/(2q)} - c_1N^{-1}}\exp\left\{-c_2\lambda^{-1/(2q)}|x-t|\right\}\right].$$

Hence, $\widetilde{f}_{ss}$ given in (3.14) is a very good approximation of $\widehat{f}_{ss}$. Results on the goodness of the approximation of $\widehat{f}_{ss}$ by $\widetilde{f}_{ss}$ are to be found in Messer (1991, theorems 3.2, 3.3), Eggermont and LaRiccia (2009, Lemma 14.7.2) and in (3.15). Summarizing the references above, the difference between $\widehat{f}_{ss}$ and $\widetilde{f}_{ss}$ at a fixed point is of a smaller order than their bias at this point and we can study asymptotic properties of $\widetilde{f}_{ss}$ instead of those of $\widehat{f}_{ss}$. From the discussion in the previous section, in the interior it holds that $\mathcal{R}_{ss}(x,t) \approx \lambda^{-1/(2q)}\mathcal{K}_{ss}\left\{\lambda^{-1/(2q)}(x-t)\right\}$. Hence, in the interior $\widehat{f}_{ss} \approx \widetilde{f}_{ss} \approx \widehat{f}_{\mathtt{K}}$, where $\widehat{f}_{\mathtt{K}}$ is a kernel estimator from (3.2) with kernel $\mathtt{K}(x) = \mathcal{K}_{ss}(x)$, bandwidth $h = \lambda^{1/(2q)}$ and $p+1 = 2q$.

The necessary moment conditions (3.3) on $\mathcal{K}_{ss}(x)$ follow from (3.23) and Messer and Goldstein (1993, Proposition 3.4):

$$\int_{-\infty}^{\infty} t^l \mathcal{K}_{ss}(t)dt \;=\begin{cases} 1, & l=0 \\ 0, & 0<l<2q \\ (-1)^{q+1}(2q)!, & l=2q. \end{cases} \tag{3.24}$$

The pointwise bias and the pointwise variance of $\widehat{f}_{\mathrm{K}}$ in the interior follow from (3.4) and from (3.5) and are given in Lemma 6.
A different approach is to be found in Nychka (1995) who worked with a basis of $W^q[0,1]$, $\{\varphi_i(x)\}_{i=1}^{\infty}$, satisfying (2.15) and its eigenvalues $\{\nu_i\}_{i=1}^{\infty}$. To obtain the pointwise bias and variance of $\widehat{f}_{ss}$, Nychka (1995) used the series representation of the Green's function from Cox (1988)

$$\mathcal{R}_\lambda(x,t) = \sum_{l=1}^{\infty} \frac{\varphi_l(x)\varphi_l(t)}{1+\lambda\nu_l}$$

and the Euler equation (3.8). The results are given for random design and for $p=1$. Restricting the design to equally spaced observations, the generalization of Nychka (1995, Theorem 2.2) for $p \geq 1$ follows directly from their proof.

Lemma 6 *(Nychka, 1995)*
Let $\widehat{f}_{ss}$ be the solution of minimization problem (2.10) with $\{x_i = i/N\}_{i=1}^N$, $f \in C^{2q}[0,1]$ and $\left|f^{(2q)}(x) - f^{(2q)}(t)\right| \leq c\,|x-t|^{\alpha}, \alpha > 0,\ 0<c<\infty$. Then for the interior of $[0,1]$ and for $c_1 N^{-2q/(2q+1)} Log(N) \leq \lambda \to 0, c_1 > 0$ it holds that

$$\begin{aligned} E\left\{\widehat{f}_{ss}(x) - f(x)\right\} &= -\,(-1)^q \lambda f^{(2q)}(x) + o(\lambda) \\ Var\left\{\widehat{f}_{ss}(x)\right\} &= \frac{\sigma^2}{N}\int_0^1 \mathcal{R}_{ss}^2(x,t)dt\,\{1+o(1)\}\,. \end{aligned}$$

Remarks

1. For $q=1$, the explicit expression for the variance is available:

$$Var\left\{\widehat{f}_{ss}(x)\right\} = \frac{\sigma^2}{8N\lambda^{1/(2)}}\,\{1+o(1)\}\,.$$

2. From Eggermont and LaRiccia (2009, p.133), the uniform bond on the smoothing spline estimator on interval $[0,1]$ is

$$\sup_{x\in[0,1]}\left|\widehat{f}_{ss}(x)-f(x)\right| = O_p\left\{\lambda^{1/2}+\sqrt{\lambda^{-1/(2q)}Log\left(N\right)/N}\right\},$$

with optimal convergence rate $\{Log(N)/N\}^{q/(2q+1)}$ when $\lambda \asymp \{Log(N)/N\}^{2q/(2q+1)}$.

3.3 Equivalent kernel for regression splines

Investigation of the local asymptotic properties of regression splines is based on different approaches. Agarwal and Studden (1980) and Zhou et al. (1998) studied the pointwise asymptotic behavior of the regression spline estimators making use of the results of Barrow and Smith (1978) on L_2-projection of a smooth function on a spline space. Another approach is to be found in Huang and Studden (1992) who derived an equivalent kernel for regression splines in terms of B-splines. We focus on the second approach since it is related to our method for finding the local asymptotic properties of penalized splines in Chapter 4.

3.3.1 Formula of the equivalent kernel for regression splines

The section is based on the result of Huang and Studden (1990) and Huang and Studden (1992). We define the equivalent kernel for regressions splines, give the explicit formula for the equivalent kernel in the interior for $p = 1, 2, 3$ and discuss its properties.
Huang and Studden (1990) defined *equivalent kernel for regression splines in the interior* $\mathcal{K}_{rs}(x,t)$ as the L_2-projection kernel on spline space $\mathcal{S}^2(p;\mathbb{Z}) = \{\sum_{i=-\infty}^{\infty} a_i B_{c,p,i}(x) : x \in \mathbb{R}, \sum_{i=-\infty}^{\infty} a_i^2 < \infty\}$, where $B_{c,p,i}(x)$ is given in (2.4). They also showed that $\mathcal{K}_{rs}(x,t)$ is the reproducing kernel of Hilbert space $\mathcal{S}^2(p;\mathbb{Z})$ with inner product $\langle f,g\rangle = \int_{-\infty}^{\infty} f(x)g(x)dx$ and found its explicit form for $p = 1, 2, 3$ in terms of B-splines:

$$\mathcal{K}_{rs}(x,t) \coloneqq \sum_{i=\lfloor x-p/2-1/2\rfloor}^{\lfloor x+p/2+1/2\rfloor}\ \sum_{j=\lfloor x-p/2-1/2\rfloor}^{\lfloor x+p/2+1/2\rfloor} B_{c,p,i}\left(x\right)B_{c,p,j}\left(t\right)\sum_{l=1}^{p} c_l\gamma_l^{|i-j|}, \tag{3.25}$$

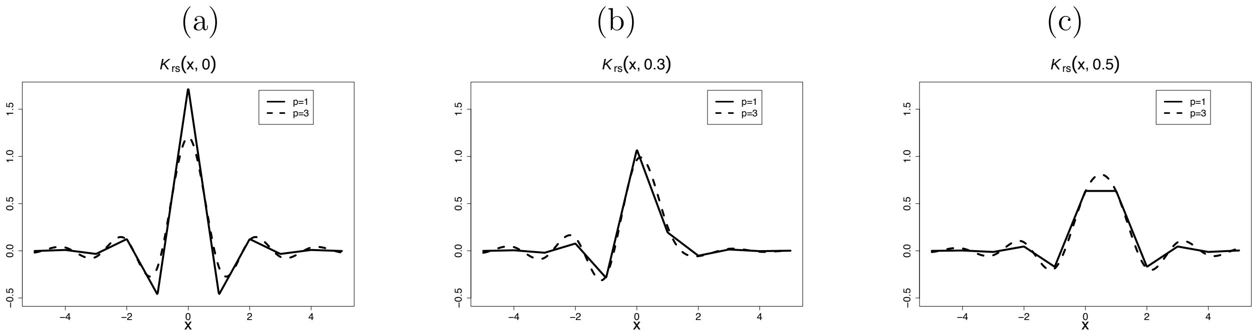

Figure 3.2: Equivalent kernel for regression splines. (a) $\mathcal{K}_{rs}(x, 0)$ for $p = 1, 3$, (b) $\mathcal{K}_{rs}(x, 0.3)$ for $p = 1, 3$, (c) $\mathcal{K}_{rs}(x, 0.5)$ for $p = 1, 3$.

where the coefficients $c_l, \gamma_l,\ l = 1, \dots, p$ are known for $p = 1, 2, 3$. For $p = 1$, $\gamma_1 = -2 + \sqrt{3}$, $c_1 = \sqrt{3}$. For $p = 2$, $\gamma_1 \simeq -0.4306$, $\gamma_2 \simeq -0.0433$, $c_1 \simeq 3.0950$, $c_2 \simeq -0.2528$. For $p = 3$, $\gamma_1 \simeq -0.5353$, $\gamma_2 \simeq -0.1226$, $\gamma_3 \simeq -0.0091$, $c_1 \simeq 6.0163$, $c_2 \simeq -1.0558$, $c_3 \simeq 0.0043$. The notation $\simeq$ is used for the approximation of the values up to the fourth place. For odd p, the formula can be written as

$$\mathcal{K}_{rs}(x, t) = \sum_{i,j=-(p+1)/2}^{(p+1)/2} B_{c,p,i}\left(\{x\}\right) B_{c,p,j}\left(\{t\}\right) \sum_{l=1}^{p} c_l \gamma_l^{|i-j+\lfloor x \rfloor - \lfloor t \rfloor|}. \tag{3.26}$$

Figure 3.2 plots $\mathcal{K}_{rs}(x, t)$ for different t. As we can see, kernel $\mathcal{K}_{rs}(x, t)$ is not translation-invariant and its shape depends on the position of t between the knots.

To obtain the equivalent kernel in form (3.25), Huang and Studden (1990) showed that the projection kernel $\mathcal{K}_{rs}(x, t)$ equals $B_\infty(x) \left\{\int_{-\infty}^{\infty} B_\infty^T(u) B_\infty(u) du\right\}^{-1} B_\infty(t)$, with $B_\infty(x) = \{B_{c,p,j}(x), j \in \mathbb{Z}\}$ being the B-spline basis vector-function of $\mathcal{S}^2(p; \mathbb{Z})$. Since $\int_{-\infty}^{\infty} B_\infty^T(u) B_\infty(u) du$ is a bi-infinite symmetric Toeplitz band matrix, its inverse matrix can be found explicitly for $p = 1, 2, 3$. The elements of the inverse matrix are of form $\left[\left\{\int_{-\infty}^{\infty} B_\infty^T(u) B_\infty(u) du\right\}^{-1}\right]_{i,j} = \sum_{l=1}^{p} c_l \gamma_l^{|i-j|},\ i, j \in \mathbb{Z}$.

Following is a list of properties of the equivalent kernel (Huang and Studden, 1990).

$$\int_{-\infty}^{\infty} \mathcal{K}_{rs}(x,t)s(t)dt = s(x),\ \forall s(x) = \sum_{i=-\infty}^{\infty} a_i B_{c,p,i}(x),\ a_i < \infty, i \in \mathbb{Z} \tag{3.27}$$

$$\int_{-\infty}^{\infty} \mathcal{K}_{rs}(x,t)t^i dt = x^i,\ i = 0,\dots,p \tag{3.28}$$

$$\int_{-\infty}^{\infty} \mathcal{K}_{rs}(x,t)t^{p+1} dt = x^{p+1} - \mathcal{B}_{p+1}(\{x\}) \tag{3.29}$$

$$|\mathcal{K}_{rs}(x,t)| \le C\gamma^{|\lfloor x\rfloor - \lfloor t\rfloor|}\ C > 0,\ 0 < \gamma < 1 \tag{3.30}$$

$$\mathcal{K}_{rs}(x,t) = \mathcal{K}_{rs}(\{x\},\{t\}). \tag{3.31}$$

Property (3.27) states that the reproducing property holds for wider class of functions than $\mathcal{S}^2(p;\mathbb{Z})$. Properties (3.28) and (3.29) give the first $p+1$ moments of $\mathcal{K}_{rs}(x,t)$ and they are obtained using (3.27) applied to polynomials $1, t, \dots, t^p$. For $i = 0$, $x = 0$ in (3.28) we use the convention $0^0 = 1$. Property (3.30) states that for fixed t kernel $\mathcal{K}(x,t)$ decays exponentially. Properties (3.28), (3.29), (3.30) are used to obtain the pointwise asymptotic behavior of regression splines.

Now we give a link between equivalent kernel $\mathcal{K}_{rs}(x,t)$ and the weight function of the regression spline estimator. Using (2.22), we write regression spline estimator $\widehat{f}_{rs}$ as

$$\widehat{f}_{rs}(x) = \frac{1}{N}\sum_{i=1}^{N} B(x)\left(N^{-1}B^TB\right)^{-1} B^T(x_i)Y_i, \tag{3.32}$$

where the design matrix B and basis vector-function $B(x)$ are given in Section 2.4. We define $\mathtt{K}_{rs}(x,t)$ as the scaled weight function

$$\mathtt{K}_{rs}(x,t) \coloneqq K^{-1}B(x/K)\left(N^{-1}B^TB\right)^{-1} B^T(t/K). \tag{3.33}$$

With this, the regression spline estimator can be written as

$$\widehat{f}_{rs}(x) = \frac{K}{N}\sum_{i=1}^{N} \mathtt{K}_{rs}(Kx, Kx_i)Y_i.$$

It appears that the equivalent kernel $\mathcal{K}_{rs}(x,t)$ is an approximation of $\mathtt{K}_{rs}(x,t)$ in the interior of interval $[0,1]$ (Huang and Studden, 1992), i.e. $\mathcal{K}_{rs}(x,t)$ is the analogue of

$\mathcal{K}_{ss}(x,t)$ for smoothing splines. A (scaled) equivalent kernel on interval $[0,1]$ is defined as

$$\mathcal{K}_{rs}^{[0,1]}(x,t) := K^{-1} B\left(K^{-1}x\right) \left\{\int_0^1 B^T(u)B(u)du\right\}^{-1} B^T\left(K^{-1}t\right). \tag{3.34}$$

Note that $\mathcal{K}_{rs}^{[0,1]}(x,t)$ corresponds to $\mathtt{K}_{rs}(x,t)$ with matrix $N^{-1}B^TB$ in (3.33) approximated by $\int_0^1 B^T(x)B(x)dx$. Huang and Studden (1992) showed that $\mathcal{K}_{rs}^{[0,1]}(x,t)$ is a good approximation of weight function $\mathtt{K}_{rs}(x,t)$ on the whole interval $[0,1]$. However, the explicit expression for $\mathcal{K}_{rs}^{[0,1]}(x,t)$ has not been obtained.

3.3.2 Pointwise asymptotic behavior of regression splines

There are two ways to find the local asymptotic properties of regression spline $\widehat{f}_{rs}$ defined in Section 2.4. The first way was considered in Huang and Studden (1992) who studied the pointwise asymptotic behavior of an equivalent kernel estimator defined as

$$\widetilde{f}_{rs}(x) := N^{-1}K\sum_{i=1}^{N} \mathcal{K}_{rs}(Kx, Kx_i)Y_i. \tag{3.35}$$

They did not consider the difference between $\widehat{f}_{rs}$ and $\widetilde{f}_{rs}$. They obtained the following bound, which gives the goodness of the approximation of $\mathtt{K}_{rs}(x,t)$ by the equivalent kernel in the interior of interval $[0,1]$:

$$|\mathtt{K}_{rs}(Kx,Kt) - \mathcal{K}_{rs}(Kx,Kt)| \leq O\left(K/N + |\gamma_1|^{K/2}\right), \ \forall t \in [0,1],$$

where γ_1 is given after formula (3.25) for $p = 1,2,3$. Together with this bound and the exponential decay of the equivalent kernel (3.30), we can conjecture that the asymptotic properties of estimators $\widehat{f}_{rs}$ and $\widetilde{f}_{rs}$ are very similar in the interior. The pointwise bias and variance of $\widetilde{f}_{rs}$ are given in Lemma 7. Although Huang and Studden (1992) considered the case $p = 3$ only, their results can easily be extended to general p.

Lemma 7 *(Huang and Studden, 1992)*
Let $f \in Lip^{p+1,\alpha}[0,1] = \{f \in C^{p+1}[0,1] : \left|f^{(p+1)}(x+\delta) - f^{(p+1)}(x)\right| \leq C\delta^\alpha$ for all $0 \leq x \leq x+\delta \leq 1\}$. Further, let $\widetilde{f}$ be the equivalent kernel estimator defined in (3.35) with equiv-

alent kernel $\mathcal{K}_{rs}(x,t)$ defined in (3.25). Let $\{x_i = i/N\}_{i=1}^N$, $K \to \infty$, $N/K \to \infty$. Then for x in the interior it holds that

$$E\left\{\widetilde{f}_{rs}(x)\right\} - f(x) = -\frac{f^{(p+1)}(x)}{(p+1)!K^{p+1}}\mathcal{B}_{p+1}(\{Kx\}) + O\left(K^{-p-1-\alpha} + \frac{1}{N}\right)$$
$$Var\left\{\widetilde{f}_{rs}(x)\right\} = \frac{\sigma^2 K}{N}\mathcal{K}_{rs}\left(\{Kx\}, \{Kx\}\right) + O\left(N^{-1}\right).$$

To obtain the pointwise bias and variance, the methods of kernel regression were used, i.e. Taylor series expansion of f and moments conditions (3.28), (3.29) of the equivalent kernel. To show that the remainder terms are asymptotically negligible, Huang and Studden (1992) used property (3.30) – the exponential decay of the kernel.
The second way of finding the local asymptotic properties of the regression spline estimator is to use the known results about the approximation error of a smooth function by splines, which was found in Barrow and Smith (1978)

$$\inf_{s\in\mathcal{S}(p-1;\underline{\tau}_K)}\left\|f(x) + \frac{f^{(p)}(\lfloor Kx\rfloor/K)}{p!K^p}\mathcal{B}_p\left(\{Kx\}\right) - s(x)\right\|_{L_\infty} = o(K^{-p}).$$

Using this result, Zhou et al. (1998) proved Lemma 8.

Lemma 8 *(Zhou et al., 1998)*
Let $f \in C^{p+1}[0,1]$, $\widehat{f}_{rs}$ be the solution of the minimization problem (2.21) with $\{x_i = i/N\}_{i=1}^N$ and $\underline{\tau}_K = \{i/K\}_{i=1}^K$. For $x \in [0,1]$,

$$E\left\{\widehat{f}_{rs}(x) - f(x)\right\} = -\frac{f^{(p+1)}(x)}{K^{p+1}(p+1)!}\mathcal{B}_{p+1}\left(\{Kx\}\right) + o(K^{-p-1})$$
$$Var\left\{\widehat{f}_{rs}(x)\right\} = \frac{\sigma^2 K}{N}\mathcal{K}_{rs}^{[0,1]}(\{Kx\}, \{Kx\}) + o(N^{-1}K).$$

With some restrictions on the design density, Lemma 8 holds for non-equidistant knots and observations.

3.4 Equivalent kernel for penalizes splines: overview

In this section, we give an overview of the related work on the pointwise asymptotic behavior of penalized spline estimator $\widehat{f}_{ps}$ defined in Section 2.5.

The first results for $p = 0, 1$, $q = 1, 2$, $k_q \to \infty$ were obtained in Li and Ruppert (2008) and for the case $p = q = 1$ are given in Lemma 9. The idea of Li and Ruppert (2008) is based on finding the inverse matrix $(B^T B + \lambda D_q)^{-1}$ in (2.24) explicitly. Since B-splines have compact support, the i-th row of matrix $(B^T B + \lambda D_q)$, except for the first and the last p rows, has form $(0, \ldots, a_p, \ldots, a_1, a_0, a_1, \ldots, a_p, \ldots, 0)$. The roots of the characteristic polynomial $a_p + a_{p-1}t + \ldots a_p t^{2p}$ are needed for the construction of the inverse matrix. For $p = 1, 2$, these roots are known and these cases were treated in Li and Ruppert (2008). For general p, we are aware of results of Xiao et al. (2011) whose idea is to approximate the roots of the characteristic polynomial for $k_q \to \infty$ by using Rouche's theorem.

A different approach and more general results can be found in Wang et al. (2011). Their idea is to search the equivalent kernel as the Green's function of certain differential equation. The authors focused on the case $p = q$. Wang et al. (2011) showed that, under certain assumptions, a penalized spline estimator is equivalent to a smoothing spline estimator.

Lemma 9 *(Wang et al., 2011)*

1. *Let* $\lambda = c^{2q} N^{-(2q)/(4q+1)}$, $c > 0$, $K \asymp N^{\tau}$, $\tau > (2q-1)/(4q+1)$, $f \in C^{2q}[0,1]$ *with* $f^{(2q)}(x) < \infty$ *and* $\hat{\gamma}(x) = O_p(K^{-1}) I_{\{p \neq q\}}$ *given in Wang et al. (2011, formulas 18,19), then for* x *in the interior*

$$N^{2q/(4q+1)} \left\{ \widehat{f}_{ps}(x) - f(x) - \hat{\gamma}(x) \right\} \xrightarrow{d} \mathcal{N} \left\{ \frac{-c^{2q}}{(-1)^q} f^{(2q)}(x), \frac{\lambda^{1/(2q)}}{c} \int_0^1 \mathcal{R}_\lambda^2(x,t) dt \right\}.$$

2. *If* $p = q$ *and* $\delta > 0$,

$$\begin{aligned} \sup_{x \in [\delta, 1-\delta]} \left| \widehat{f}_{ss}(x) - \widehat{f}_{ps}(x) \right| =& O_p\left(\frac{\lambda^{1/2}}{K} \right) + O_p \left\{ \sqrt{\frac{Log(K)}{N\lambda K}} \right\}, \\ \sup_{x \in [0,1]} \left| \widehat{f}_{ss}(x) - \widehat{f}_{ps}(x) \right| =& O_p\left(\frac{1}{K} \right) + O_p \left(\sqrt{\frac{Log(K)}{N\lambda K}} \right). \end{aligned}$$

Note, that the first part of Lemma 9 corresponds to $k_q \to \infty$ scenario, since $k_q \asymp \lambda^{1/2q} K \asymp N^{(2q-2)/(4q+1)} N^{\xi}$, $0 < \xi \le (2q+2)/(4q+1)$. We recall that the optimal convergence rate in the interior for $k_q \to \infty$ is $N^{-2q/(4q+1)}$. Hence, for term $\hat{\gamma}(x) =$

$O_p(N^{\tau})I_{\{p\neq q\}}$ being asymptotically negligible we need, in fact, $\tau > 2q/(4q+1)$ for $p \neq q$. The bounds in the second part of the Lemma 9 are sufficient only for the $k_q \to \infty$ scenario, since for $k_q = O(1)$, the bounds are of the bigger order then the bias, which is $O(K^{-2q})$. For $k_q \to \infty$, the convergence rate of the penalized spline estimator to the smoothing spline estimator, applied for $\lambda \asymp N^{-2q/(4q+1)}$ and $K \asymp N^{\tau}$, $\tau > (2q-1)/(4q+1)$, is $N^{-\xi}Log(N)$, $\xi > 2q/(4q+1)$ in the interior. The convergence rate at the boundary is slower.

Summarizing, the available results on the pointwise asymptotic behavior of penalized splines are restrictive and only the case $k_q \to \infty$ was considered. In Section 4.3 and Section 4.4, we give the local and the global asymptotic properties of penalized splines for different k_q-scenarios in the periodic case.

4 General framework for splines

In this chapter, our main results are stated and proved. In the special case of periodic functions, we find the Demmler-Reinsch basis, the equivalent kernel for periodic spline estimators, and study the global and the local asymptotic properties of these estimators. Thereby we investigate smoothing, regression and penalized splines simultaneously.
As before, we consider regression model (2.1) with standard error assumptions (2.2). Assume that the unknown regression function f is periodic with period 1, more precisely,

$$f \in C^{p+1}_{\mathtt{per}} = \{f : \ f \in C^{p+1}(\mathbb{R}), \ f^{(j)}(0+l) = f^{(j)}(1+l), \ l \in \mathbb{Z}, \ j = 0, \ldots, p\}.$$

Assume that the observations points $\{x_i = i/N\}_{i=1}^{N}$ are fixed and equally spaced on $[0,1]$ interval. *Periodic penalized spline* estimator $\widehat{f}$ of f is defined as the minimizer of

$$\min_{s \in \mathcal{S}_{\mathtt{per}}(p;\underline{\tau}_K)} \left[\frac{1}{N} \sum_{i=1}^{N} \{Y_i - s(x_i)\}^2 + \lambda \int_0^1 \left\{ s^{(q)}(x) \right\}^2 dx \right], \ \lambda > 0, \ 0 < q \le p, \tag{4.1}$$

where $\mathcal{S}_{\mathtt{per}}(p;\underline{\tau}_K)$ is a space of periodic splines of degree p based on equidistant knots $\underline{\tau}_K = \{i/K\}_{i=0}^{K}$ (see Definition 4). This chapter studies the behavior of $\widehat{f}$. Without loss of generality, assume number of knots K to be even and $M = N/K$ – the number of observations in each interval $[\tau_i, \tau_{i+1})$ – to be an integer.
Periodic penalized spline estimator $\widehat{f}$ can occur in two extreme cases. If number of knots K equals number of observations N and $\lambda > 0$, estimator $\widehat{f}$ is a periodic smoothing spline estimator (see Section 2.3). If penalty parameter $\lambda = 0$ and $K \ll N$, then $\widehat{f}$ is a periodic regression spline estimator (see Section 2.4). From now on, we will refer to $\widehat{f}$ as the periodic spline estimator, stressing that all cases are included.
In Section 4.1, we introduce the common Demmler-Reinsch basis for all periodic spline estimators. In Section 4.2, we study the global asymptotics (the IMSE) of periodic spline

estimators. In Section 4.3, we find a bandwidth that is universal for all splines and study weight functions of periodic spline estimators. In Section 4.4, we use the results of Section 4.3 to find the pointwise bias and variance of periodic spline estimators.

4.1 Demmler-Reinsch basis for periodic splines

We recall that periodic cardinal B-splines $B_{i,p}(x), i = 1, \ldots, K$ given in (2.6) build a basis of $\mathcal{S}_{\texttt{per}}(p; \underline{\tau}_K)$. This basis has several useful properties, including compact support (see Section 2.1.2) and known Fourier series representation (2.7). However, the basis is not orthonormal and, hence, not convenient for studying periodic spline estimators. In this section, we will give the exact expression for the orthonormal Demmler-Reinsch basis of periodic splines in terms of exponential splines. Similar to natural spline space $\mathcal{NS}(2q-1; \underline{\tau}_{N+1})$ in Section 2.3, the exponential Demmler-Reinsch basis $\{\phi_i(x)\}_{i=1}^K$ of periodic spline space $\mathcal{S}_{\texttt{per}}(p; \underline{\tau}_K)$ is determined by conditions

$$\frac{1}{N}\sum_{l=1}^{N}\phi_i(x_l)\overline{\phi_j(x_l)} = \delta_{i,j}; \qquad \int_0^1 \phi_i^{(q)}(x)\overline{\phi_j^{(q)}(x)}dx = \delta_{i,j}\mu_i; \qquad \phi_j(x) = \overline{\phi_{K-i}(x)}. \tag{4.2}$$

The construction of the Demmler-Reinsch basis is performed in three steps. First, in Section 4.1.1, we recall the definition of known exponential splines, Euler-Frobenius and Q-polynomials, study their properties and connections. Most properties are known and given without proofs, but some properties needed in subsequent proves are derived. Second, in Section 4.1.2, we construct special trigonometric polynomials $Q_{p,M}$. These polynomials depend on the ratio $M = N/K$ and arise in case the number of knots and observations are different. Third, in Section 4.1.3, we define the Demmler-Reinsch basis via exponential splines and $Q_{p,M}$-polynomials and verify that the basis satisfies all necessary conditions using the properties stated in Section 4.1.1. Finally, we find coefficients of periodic spline estimator $\widehat{f}$ in the Demmler-Reinsch and in the Fourier basis.

4.1.1 Euler-Frobenius polynomials, exponential splines and Q-polynomials

The section is organized as follows. First, we give a definition of well-known Euler-Frobenius polynomials and list their properties. Second, we define more general functions that were studied intensively in Schoenberg (1973) – exponential splines. We list properties of exponential splines we will need later. Finally, we define Q-polynomials that are related to Euler-Frobenius polynomials and and are used intensively in this dissertation. We give a brief literature overview, since Q-polynomials (or very similar polynomials) often arise in different contexts, although do not have common notation.

Definition 14 *Euler-Frobenius polynomials are defined via the recurrence relation*

$$\begin{aligned} \Pi_{n+1}(t) &:= (1+nt)\Pi_n(t) + t(1-t)\Pi_n^{'}(t) \\ \Pi_0(t) &:= 1. \end{aligned} \tag{4.3}$$

The first five Euler-Frobenius polynomials are

$$\begin{aligned} \Pi_1(t) &= 1 \\ \Pi_2(t) &= t+1 \\ \Pi_3(t) &= t^2+4t+1 \\ \Pi_4(t) &= t^3+11t^2+11t+1 \\ \Pi_5(t) &= t^4+26t^3+66t^2+26t+1. \end{aligned}$$

Following is a list of properties of the Euler-Frobenius polynomials.

1. The connection with B-splines (He, 2011)

$$\Pi_p(t) = p! \sum_{j=0}^{p-1} B_{c,p}(p-j)t^j.$$

2. The explicit formula for the Euler-Frobenius polynomials

$$\Pi_p(t) = \sum_{i=0}^{p} \binom{p+1}{i} (-1)^i \sum_{j=1}^{p-i} j^p t^{p-j-i},\ p>0, \text{with } 0^0 := 1.$$

3. Roots of$\Pi_p(t)$ are simple and negative. Denoting them $\lambda_{p-1} < \lambda_{p-2} < \ldots < \lambda_1 < 0$, it holds that $\lambda_1\lambda_{p-1} = \lambda_2\lambda_{p-2}\ldots = 1$ (Schoenberg, 1973).

Using Euler-Frobenius polynomials, we can define exponential splines.

Definition 15 *Exponential spline of degree p to base z are defined as follows.*

$$\Phi_p(t,z) := z^{\lfloor t \rfloor}\left(1 - z^{-1}\right)^p \sum_{j=0}^{p} \binom{p}{j} \frac{\{t\}^{p-j}\,\Pi_j(z)}{p!\,(z-1)^j},\ z \neq 0,\ z \neq 1, \tag{4.4}$$

With the convention $0^0 := 1$*, we define* $\Phi_p(t,z) := z^{t-p}\Pi_p(z)/p!, t \in \mathbb{Z}$ *and* $\Phi_p(t,1) := 1$.

The following properties hold for exponential splines.

1. Schoenberg (1973, Lecture 2, Lemma 2)

$$\Phi_p(t+1,z) = z\Phi_p(t,z)\,, \text{ on the interval } 0 < t < 1. \tag{4.5}$$

2. An equivalent definition of the exponential splines (Schoenberg, 1973, Lecture 2)

$$\Phi_p(t,z) = \sum_{l=-\infty}^{\infty} z^l B_{c,p}(t-l)\,,\ z \neq 0,\ z \neq 1.$$

3. Fourier series representation

$$\Phi_p\{t, \exp(2\pi \mathtt{i} z)\} = \frac{\exp(2\pi \mathtt{i} zt)}{\exp\{\pi \mathtt{i} z(p+1)\}} \sum_{l=-\infty}^{\infty} \frac{\operatorname{sinc}\{\pi(z+l)\}^{p+1}\exp(2\pi \mathtt{i} lt)}{(-1)^{l(p+1)}}. \tag{4.6}$$

The last property needs to be derived.

Proof of (4.6)

The proof largely follows from Theorem 5 in Lecture 3 of Schoenberg (1973). From formulas (1.1) and (1.4) given in the lecture follows the equality

$$\frac{\exp(2\pi \mathtt{i} z) - 1}{\exp(2\pi \mathtt{i} z) - \exp(x)} \frac{\exp(\{t\}x)}{x^{p+1}} = \sum_{l=0}^{\infty} \frac{\{1 - \exp(-2\pi \mathtt{i} z)\}^{-l}\,\Phi_l\{t, \exp(2\pi \mathtt{i} z)\}}{\exp(2\pi \mathtt{i} z\,\lfloor t \rfloor)} x^{l-p-1}.$$

The residue of this function at 0 is $\exp(-2\pi \mathtt{i} z\,\lfloor t \rfloor)\{1 - \exp(-2\pi \mathtt{i} z)\}^{-p}\,\Phi_p\{t, \exp(2\pi \mathtt{i} z)\}$, while $\{\exp(-2\pi \mathtt{i} z) - 1\}\exp\{2\pi \mathtt{i}\{t\}(z+l)\}/\{2\pi \mathtt{i}(z+l)\}^{p+1}$ are the residues at poles

$2\pi\mathtt{i}\,(z+l)\,,\ l \in \mathbb{Z}$. With this, from the Cauchy residue theorem, it follows that

$$\sum_{l=-\infty}^{\infty} \frac{\{1-\exp(-2\pi\mathtt{i}z)\}\exp\{2\pi\mathtt{i}\{t\}(z+l)\}}{\{2\pi\mathtt{i}\,(z+l)\}^{p+1}} = \frac{\{1-\exp(-2\pi\mathtt{i}z)\}^{-p}\,\Phi_p\{t,\exp(2\pi\mathtt{i}z)\}}{\exp(2\pi\mathtt{i}z\,\lfloor t\rfloor)}.$$

Multiplication by $\sin(\pi z)^{p+1}$ and some simplifications lead to

$$\Phi_p\{t,\exp(2\pi\mathtt{i}z)\} = \frac{\{1-\exp(-2\pi\mathtt{i}z)\}^{p+1}\exp(2\pi\mathtt{i}zt)}{\sin(\pi z)^{p+1}\,(2\mathtt{i})^{p+1}} \sum_{l=-\infty}^{\infty} \frac{\sin\{\pi(z+l)\}^{p+1}\exp(2\pi\mathtt{i}tl)}{(-1)^{l(p+1)}\{\pi(z+l)\}^{p+1}}.$$

Making use of $\sin(\pi x) = \{\exp(\pi\mathtt{i}x) - \exp(-\pi\mathtt{i}x)\}/(2\mathtt{i})$ proves (4.6). □

Finally, we define Q-polynomials as series and will show later in Lemma 10 that they are, in fact, trigonometric polynomials.

Definition 16 *Q-polynomials of degree $p-1$ are defined as series*

$$Q_{p-1}(z) \coloneqq \sum_{l=-\infty}^{\infty} \operatorname{sinc}\{\pi(z+l)\}^{p+1}. \tag{4.7}$$

In a somewhat different context, Q-polynomials have been discussed in Gautschi (1971), who studied the so-called *attenuation factors* in the approximation of the Fourier coefficients of f available on a grid of N values $\{f(x_l)\}_{l=1}^N$. In the simplest case, the i-th attenuation factor is defined as the ratio of i-th Fourier coefficient of an interpolation of values $\{f(x_l)\}_{l=1}^N$ and $N^{-1}\sum_{l=1}^N f(x_l)\exp(-2\pi\mathtt{i}ix_l)$. If we interpolate values $\{f(x_l)\}_{l=1}^N$ by a periodic spline of degree $2q-1$, then the i-th attenuation factor is $\operatorname{sinc}(\pi i/N)^{2q}/Q_{2q-2}(i/N)$ (see Example 5.1 in Gautschi (1971) and σ-and q-polynomials defined in his Section 2). Q-polynomials also arise in Blu and Unser (1999) as an interpolation filter for B-splines and in spline wavelet analysis (e.g., Zheludev, 1996).

The next lemma states that $Q_{p-1}(z)$ are polynomial of $\cos(\pi z)$ of degree $(p-1)$ and gives a recursive formula for them.

Lemma 10 *Polynomial $Q_{p-1}(z)$ defined in (4.7) can be obtained using the following* *recursive formula*

$$Q_j(z) = \cos(\pi z)Q_{j-1}(z) + \frac{1-\cos(\pi z)^2}{j+1}\frac{dQ_{j-1}(z)}{d\cos(\pi z)},\ j = 1,\ldots,p-1, z \notin \mathbb{Z}, \qquad (4.8)$$

where for odd p*, the recursive formula (4.8) is applied with* $Q_0(z) := 1$*, while for even* p*, formula (4.8) is applied with* $Q_0(z) := \cos(\pi z)$*. If* $z \in \mathbb{Z}$*,* $Q_{p-1}(z) = 1$*.*

Proof of lemma 10

For $z \in \mathbb{Z}$ equality $Q_{p-1}(z) = 1$ follows immediately from (4.7) . We consider case $z \notin \mathbb{Z}$. For odd p, series (4.7) can be written as

$$\sum_{l=-\infty}^{\infty} \mathrm{sinc}\{\pi(z+l)\}^{p+1} = \frac{\sin(\pi z)^{p+1}}{\pi^{p+1}} \sum_{l=-\infty}^{\infty} (z+l)^{-(p+1)} = (-1)^p \frac{\sin(\pi z)^{p+1}}{p!\pi^{p+1}} \frac{\partial^p}{\partial z^p} \sum_{l=-\infty}^{\infty} (z+l)^{-1}.$$

For even p, it holds that

$$\sum_{l=-\infty}^{\infty} \mathrm{sinc}\{\pi(z+l)\}^{p+1} = \frac{\sin(\pi z)^{p+1}}{\pi^{p+1}} \sum_{l=-\infty}^{\infty} \frac{(z+l)^{-(p+1)}}{(-1)^l} = (-1)^p \frac{\sin(\pi z)^{p}}{p!\pi^{p}} \frac{\partial^p}{\partial z^p} \sum_{l=-\infty}^{\infty} \frac{(z+l)^{-1}}{(-1)^l}.$$

It is known that $\sum_{l=-\infty}^{\infty}(z+l)^{-1} = \pi\cot(\pi z)$ and $\sum_{l=-\infty}^{\infty}(-1)^l(z+l)^{-1} = \pi\sin(\pi z)^{-1}$, where the summation is understood in the principal value sense. Hence, for odd p, the functions in recursion (4.8) are

$$Q_{j-1}(z) := (-1)^j \frac{\sin(\pi z)^{j+1}}{j!\pi^j} \frac{\partial^j}{\partial z^j} \cot(\pi z),\ j = 1,\ldots,p.$$

For even p, these functions are

$$Q_{j-1}(z) := (-1)^j \frac{\sin(\pi z)^{j+1}}{j!\pi^j} \frac{\partial^j}{\partial z^j} \sin(\pi z)^{-1},\ j = 1,\ldots,p.$$

Substitution of the corresponding functions in (4.8) proves the lemma. □

Applying Lemma 10, we obtain the first three $Q_{p-1}(z)$ polynomials for odd p

$$\begin{aligned}
Q_2(z) &= 1/3 + 2\cos(\pi z)^2/3 \\
Q_4(z) &= 2/15 + 11\cos(\pi z)^2/15 + 2\cos(\pi z)^4/15 \\
Q_6(z) &= 17/315 + 4\cos(\pi z)^2/7 + 38\cos(\pi z)^4/105 + 4\cos(\pi z)^6/315.
\end{aligned}$$

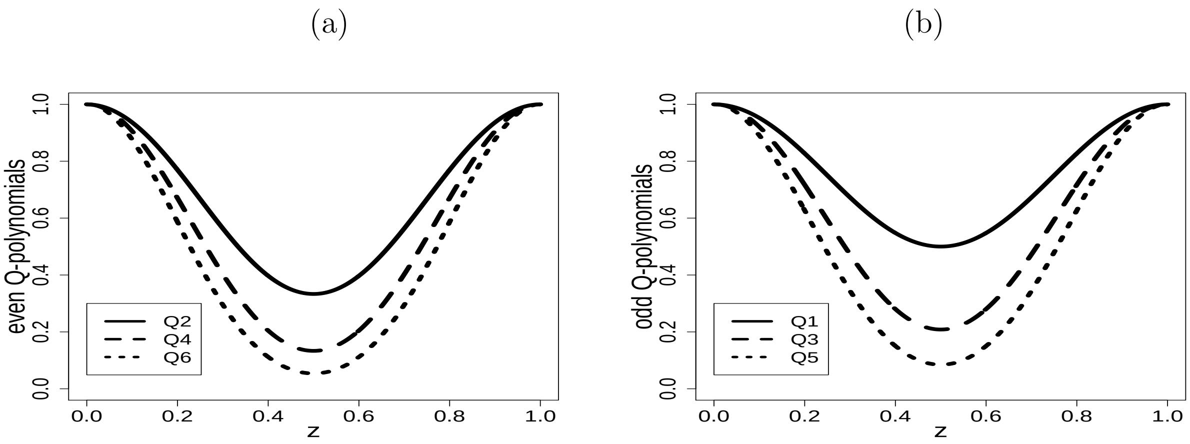

Figure 4.1: (a) Q_{p-1} polynomials for p odd, (b) Q_{p-1} polynomials for p even

For even p, the first three $Q_{p-1}(z)$ polynomials are given by

$$\begin{aligned}
Q_1(z) &= 1/2 + \cos(\pi z)^2/2 \\
Q_3(z) &= 5/24 + 3\cos(\pi z)^2/4 + \cos(\pi z)^4/24 \\
Q_5(z) &= 61/720 + 479\cos(\pi z)^2/720 + 179\cos(\pi z)^4/720 + \cos(\pi z)^6/720.
\end{aligned}$$

Later we will make use of the following connection between the Euler-Frobenious and Q-polynomials.

Lemma 11 *Let $\Pi_p(t)$ be the Euler-Frobenius polynomial defined in (4.3). For odd p, Q-polynomials can be expressed as*

$$Q_{p-1}(z) = \exp\{\mathtt{i}z\pi(p-1)\}\Pi_p\{\exp(-2\mathtt{i}\pi z)\}/p!. \tag{4.9}$$

For even p, Q-polynomials can be expressed as

$$\begin{aligned}
Q_{p-1}(z) &= \exp\{\pi\mathtt{i}z(p-1)/2\}\cos(\pi z/2)^{p+1}\,\Pi_p\{\exp(-\pi\mathtt{i}z)\}/p! \\
&- (-1)^{p/2}\mathtt{i}\exp\{\pi\mathtt{i}z(p-1)/2\}\sin(\pi z/2)^{p+1}\,\Pi_p\{-\exp(-\pi i z)\}/p!.
\end{aligned}$$

Proof of lemma 11
We apply formulas (4.4) and (4.6) for$\Phi\ _p\{0, \exp(-2\pi \mathtt{i} z)\}$ to obtain

$$\exp\left\{\pi \mathtt{i} z\left(p-1\right)\right\}\Pi_p\left\{\exp\left(-2\pi \mathtt{i} z\right)\right\}/p! = \sum_{l=-\infty}^{\infty}\left(-1\right)^{l(p+1)}\operatorname{sinc}\left\{\pi\left(z+l\right)\right\}^{p+1}. \quad (4.10)$$

With this, the lemma is proved for odd p. For even p, we use (4.10) and the fact that

$$\begin{aligned}\sum_{l=-\infty}^{\infty}\operatorname{sinc}\left\{\pi\left(z+l\right)\right\}^{p+1} &= \cos\left(\pi z/2\right)^{p+1}\sum_{l=-\infty}^{\infty}(-1)^l\operatorname{sinc}\left\{\pi\left(z/2+l\right)\right\}^{p+1}\\ &\quad -\sin\left(\pi z/2\right)^{p+1}\sum_{l=-\infty}^{\infty}(-1)^l\operatorname{sinc}\left\{\pi(z+1)/2+\pi l\right\}^{p+1}.\end{aligned}$$

□

4.1.2 Q-polynomials for low-rank splines

As shown in (4.2), the Demmler-Reinsch basis is an orthonormal basis of a spline space under discrete inner product $\langle s_1, s_2\rangle = N^{-1}\sum_{l=1}^{N} s_1(x_l)s_2(x_l)$, where s_1, s_2 are spline functions from a corresponding spline space and $\{x_l\}_{l=1}^{N}$ are the observation points. Polynomials $Q_{p,M}$, which we define in this section, are used for the normalization of the Demmler-Reinsch basis for periodic splines under the discrete inner product and appear in special case $K \neq N$. For $K = N$, they can be expressed in terms of Q-polynomials. We define

$$\begin{aligned}Q_{p,M}\left(z\right) &:= \frac{1}{N}\sum_{i=1}^{N}\left|\Phi_p\left\{K(i/N)+(p+1)/2, \exp(-2\pi \mathtt{i} z)\right\}\right|^2\\ &= \frac{1}{N}\sum_{i=1}^{M}\sum_{j=0}^{K-1}\left|\Phi_p\left\{(i+jM)/M+\left(p+1\right)/2, \exp(-2\pi \mathtt{i} z)\right\}\right|^2\\ &= \frac{1}{M}\sum_{i=1}^{M}\left|\Phi_p\left\{i/M+(p+1)/2, \exp(-2\pi \mathtt{i} z)\right\}\right|^2, \quad (4.11)\end{aligned}$$

where on the last step we used $N = MK$ and the fact that

$$|\Phi_p\{t+j, \exp(-2\pi\mathtt{i}z)\}| = |\Phi_p\{t, \exp(-2\pi\mathtt{i}z)\}|\,, j \in \mathbb{Z},$$

which follows from (4.5). Let us examine the properties of $Q_{p,M}$-polynomials.
From (4.11), functions $Q_{p,M}$ varies smoothly between $Q_{2p}(z)$ and $Q_{p-1}^2(z)$, depending on M. For $M = 1$, we find from (4.6) that $|\Phi_p\{i + (p+1)/2, \exp(-2\pi\mathtt{i}z)\}| = Q_{p-1}(z)$. Therefore, it holds that

$$Q_{p,1}(z) = Q_{p-1}^2(z).$$

For $M = N/K \to \infty$, we approximate (4.11) by an integral and use series representation (4.6) of the exponential splines, obtaining

$$\begin{aligned}
\lim_{M\to\infty} Q_{p,M}(z) &= \int_0^1 |\Phi_p\{x + (p+1)/2, \exp(-2\pi\mathtt{i}z)\}|^2\, dx \\
&= \int_0^1 \sum_{l,s=-\infty}^{\infty} \operatorname{sinc}\{\pi(z+l)\}^{p+1} \operatorname{sinc}\{\pi(z+s)\}^{p+1} \exp\{2\pi\mathtt{i}(l-s)x\}\, dx \\
&= \sum_{l=-\infty}^{\infty} \operatorname{sinc}\{\pi(z+l)\}^{2p+2} = Q_{2p}(z).
\end{aligned}$$

Lemma 12 gives the connection between $Q_{p,M}$ and the Euler-Frobenius polynomials and, with this, an explicit formula for $Q_{p,M}$.

Lemma 12 *We define function $\widetilde{\Pi}_{p,M}$ as*

$$\widetilde{\Pi}_{p,M}(u) \coloneqq \sum_{j,l=0}^{p} \frac{\Pi_j(u)\Pi_l(u^{-1})u^l(u-1)^{2p-l-j}}{(-1)^{p-l} j!l!\,(p-j)!(p-l)!} \sum_{s=0}^{2p-l-j} \binom{2p-l-j}{s} 2^s \frac{\mathcal{B}_{s+1}(M) - \mathcal{B}_{s+1}}{(s+1)(2M)^{2p-j-l+1}},$$

if for p even and M odd. In all other cases

$$\widetilde{\Pi}_{p,M}(u) \coloneqq \sum_{j,l=0}^{p} \frac{\Pi_j(u)\Pi_l(u^{-1})\, u^l(u-1)^{2p-l-j}}{(-1)^{p-l} j!l!\,(p-j)!\,(p-l)!} \frac{\mathcal{B}_{2p-j-l+1}(M) - \mathcal{B}_{2p-j-l+1}}{M^{2p-j-l+1}\,(2p-j-l+1)}.$$

Additionally we define $\widetilde{\Pi}_{p,M}(1) \coloneqq \lim_{u\to 1} \widetilde{\Pi}_{p,M}(u) = 1$. Then $Q_{p,M}$-polynomials can be *expressed in terms of* $\widetilde{\Pi}_{p,M}$*-polynomials by*

$$Q_{p,M}(z) = \exp(2\pi \mathtt{i} pz)\widetilde{\Pi}_{p,M}\{\exp(-2\pi \mathtt{i} z)\}.$$

Proof of lemma (12)
Using (4.4) and (4.11), we compute $N^{-1}\sum_{i=1}^{N}|\Phi_p(\{i/M+(p+1)/2,u\})|^2$ with $|u| = 1, u \neq 1$:

$$\begin{aligned}&\sum_{i=1}^{N}\frac{(1-u)^p(1-u^{-1})^p}{N}\sum_{j,l=0}^{p}\binom{p}{j}\binom{p}{l}\frac{\{i/M+(p+1)/2\}^{2p-j-l}\,\Pi_j(u)\Pi_l(u^{-1})}{(p!)^2\,(u-1)^j(u^{-1}-1)^l}\\ =&\frac{1}{M}\sum_{j,l=0}^{p}\frac{\Pi_j(u)\Pi_l(u^{-1})(u-1)^{2p-j-l}}{(-1)^{p+l}u^{p-l}(p-j)!(p-l)!}\sum_{i=1}^{M}\{i/M+(p+1)/2\}^{2p-j-l}.\end{aligned}$$

Using the property of Bernoulli polynomials $\sum_{i=0}^{M-1} i^{j-1} = \{\mathcal{B}_j(M) - \mathcal{B}_j\}/j$ and some easy simplifications prove the lemma. □

From Lemma 11 and 12, it also follows that functions $Q_{p,M}$ are indeed trigonometrical polynomials, justifying our vocabulary.

Using Lemma 12, we find $Q_{p,M}$-polynomials for $p = 1,\ldots,5$.

$$\begin{aligned}
Q_{1,M}(z) &= Q_2(z) + \frac{2\sin(\pi z)^2}{3M^2}\\
Q_{2,M}(z) &= \begin{cases} Q_4(z) - \frac{2\sin(\pi z)^4}{15M^4}, & \text{for even M}\\ Q_4(z) + \frac{7\sin(\pi z)^4}{60M^4}, & \text{for odd M}\end{cases}\\
Q_{3,M}(z) &= Q_6(z) + 2\frac{3\sin(\pi z)^4 - 2\sin(\pi z)^6}{135M^4} + \frac{8\sin(\pi z)^6}{189M^6}\\
Q_{4,M}(z) &= \begin{cases} Q_8(z) - \frac{2\sin(\pi z)^8}{135M^8} - \frac{4\{2+\cos(2\pi z)\}\sin(\pi z)^6}{567M^6}, & \text{for even M}\\ Q_8(z) + \frac{127\sin(\pi z)^8}{8640M^8} + \frac{31\{2+\cos(2\pi z)\}\sin(\pi z)^6}{4536M^6}, & \text{for odd M}\end{cases}\\
Q_{5,M}(z) &= Q_{10}(z) + \frac{2\{2+\cos(2\pi z)\}}{675M^8\sin(\pi z)^{-8}} + \frac{33+26\cos(2\pi z)+\cos(4\pi z)}{14175M^6\sin(\pi z)^{-6}} + \frac{8\sin(\pi z)^{10}}{1485M^{10}}.
\end{aligned}$$

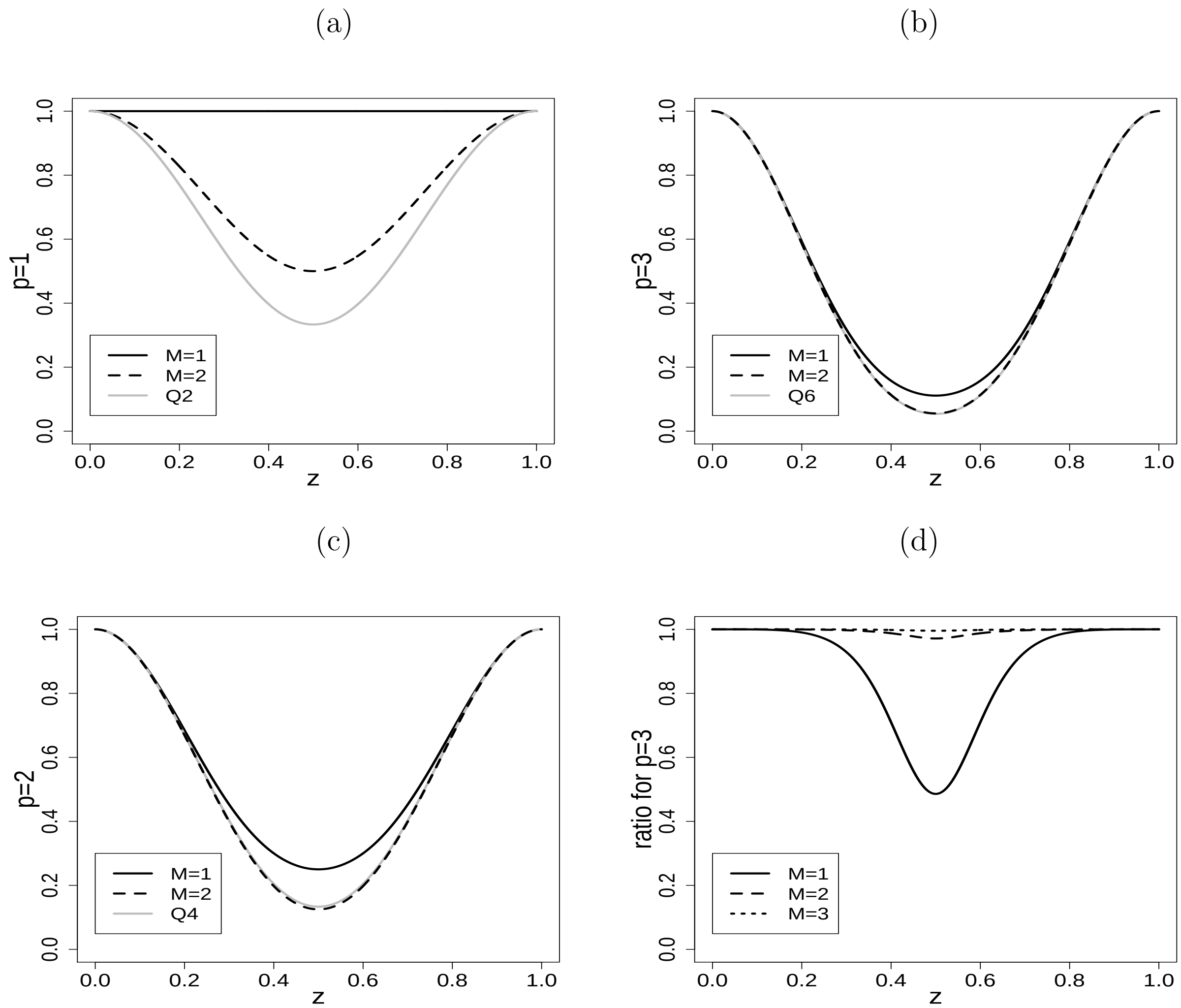

Figure 4.2: Functions $Q_{p,M}(z)$: (a) $Q_{1,M}$ for $M = 1, 2, 3$, (b) $Q_{3,M}$ for $M = 1, 2$, (c) $Q_{2,M}$ for $M = 1, 2$, (d) $Q_6/Q_{p,M}$ for different M. Grey line denotes Q_{2p}.

Plots of $Q_{p,M}$-polynomials are given in Figure 4.1. In (a), (b), (c), the grey lines correspond to Q_{2p}-polynomials for different p, while the solid black lines correspond to squared Q_{p-1}-polynomials. In (a), (b), we can see how $Q_{p,M}$-polynomials for odd p change between $Q_{2p}(z)$ and $Q_{p-1}^2(z)$ depending on M. For $p = 3$, this change is already so rapid that $Q_6(z)$ and $Q_{3,2}(z)$ practically coincide and only in (d) we can see that they are slightly different in the interior. In (c), we see that the $Q_{p,M}(z)$ are bounded by $Q_{p,2}(z)$ and $Q_{p-1}^2(z)$ for even p and converge very fast to $Q_{2p}(z)$ with growing M.

A few more properties of $Q_{p,M}$-polynomials, such as a series representation and the connection to Q-polynomials, are given below.

1. Series representation

$$Q_{p,M}(z) = \sum_{l,j=-\infty}^{\infty} \operatorname{sinc}\left\{\pi(z+l)\right\}^{p+1} \operatorname{sinc}\left\{\pi(z+l+jM)\right\}^{p+1}. \tag{4.12}$$

2. Connection to Q_{2p}-polynomials

$$Q_{p,M}(z) = Q_{2p}(z) + \sum_{i=0}^{M-1}\sum_{l\neq j}\left[\operatorname{sinc}\left\{\pi(z+i+lM)\right\}\operatorname{sinc}\left\{\pi(z+i+jM)\right\}\right]^{p+1}.$$

3. Bounds

$$0 < Q_{2p}(z) \leq Q_{p,M}(z) \leq Q_{p-1}^2(z), \text{ if } p \text{ or } M \text{ are odd}$$
$$0 < Q_{p,2}(z) \leq Q_{p,M}(z) \leq Q_{2p}(z), \text{ if } p \text{ even and } M \text{ are even.}$$

4. Bounds for ratio $Q_{2p}(z)/Q_{p,M}(z)$

$$Q_{2p}(z)Q_{p-1}^{-2}(z) \leq Q_{2p}(z)Q_{p,M}^{-1}(z) \leq 1, \text{ if } p \text{ or } M \text{ are odd}$$
$$1 \leq Q_{2p}(z)Q_{p,M}^{-1}(z) \leq Q_{2p}(z)Q_{p,2}^{-1}(z) < 3/2, \text{ if } p \text{ even and } M \text{ are even.}$$

4.1.3 Demmler-Reinsch basis and Fourier coefficients of spline estimators

Even though the Demmler-Reinsch basis for periodic smoothing splines has been employed in Cogburn and Davis (1974) and Craven and Wahba (1979), no explicit expressions were derived there. Our next lemma gives the explicit expression for the complex-valued Demmler-Reinsch basis for periodic spline space $S_{\mathtt{per}}(p; \underline{\tau}_K)$.

Lemma 13 *For $x \in \mathbb{R}$, functions*

$$\phi_i\,(x) \coloneqq \frac{\Phi_p\{Kx + (p+1)/2, \exp(-2\pi\mathtt{i}i/K)\}}{\sqrt{Q_{p,M}(i/K)}} \tag{4.13}$$

$i = 1, \ldots, K$ form the complex-valued Demmler-Reinsch basis in $S_{\mathtt{per}}(p; \underline{\tau}_K)$, i.e.

$$\begin{aligned} \frac{1}{N}\sum_{l=1}^{N} \phi_i(l/N)\overline{\phi_j(l/N)} &= \delta_{i,j} \qquad (4.14)\\ \int_0^1 \phi_i^{(q)}(x)\overline{\phi_j^{(q)}(x)}dx &= \mu_i\delta_{i,j}, \qquad (4.15) \end{aligned}$$

i,j=1,…, K, where $\delta_{i,j}$ is the Kronecker's delta and the eigenvalues are

$$\mu_i \coloneqq (2\pi i)^{2q}\mathrm{sinc}(\pi i/K)^{2q}\frac{Q_{2p-2q}(i/K)}{Q_{p,M}(i/K)}. \tag{4.16}$$

Proof of Lemma 13

The proof is based on the series representation of the Demmler-Reinsch basis and its connection to the discrete Fourier transform (DFT) of periodic B-splines. Using (4.6) and the definition of the complex-valued Demmler-Reinsch basis (4.13), we can write

$$\phi_i(x) = \frac{1}{\sqrt{Q_{p,M}(i/K)}}\sum_{l=-\infty}^{\infty} \mathrm{sinc}\{\pi(i/K + l)\}^{p+1}\exp\{-2\pi\mathtt{i}x(i + lK)\}. \tag{4.17}$$

Plugging the Fourier series of a periodic B-spline (2.7) into the DFT of B-splines, we find

$$\begin{aligned} \sum_{i=1}^{K} B_{i,p}(x)\exp(-2\pi\mathtt{i}li/K) &= \sum_{m=-\infty}^{\infty} \mathrm{sinc}(\pi m/K)^{p+1}\exp(-2\pi\mathtt{i}mx)\\ &\times \sum_{i=1}^{K}\exp\{2\pi\mathtt{i}i(m-l)/K\}\\ &= K\sum_{n=-\infty}^{\infty}\mathrm{sinc}\{\pi(l/K+n)\}^{p+1}\exp\{-2\pi\mathtt{i}(l+nK)x\}\\ &= K\sqrt{Q_{p,M}(l/K)}\phi_l(x), \end{aligned}$$

where in the last equality representation (4.17) has been used and $n = (m-l)/K$. The properties of DFT ensure that the functions $\phi_i(x)$, $i = 1, \ldots, K$ are also the basis in $S_{\mathtt{per}}(p; \underline{\tau}_K)$. Moreover, if the $\{\gamma_i\}_{i=1}^K$ and $\{\beta_i\}_{i=1}^K$ are the coefficients of spline function $s \in S_{\mathtt{per}}(p; \underline{\tau}_K)$ in different bases, i.e. $s(x) = \sum_{i=1}^K B_{i,p}(x)\gamma_i = \sum_{i=1}^K \phi_i(x)\beta_i$, then $\gamma_i = \sum_{l=1}^K \beta_l \exp(-2\pi \mathtt{i} li/K)/\{KQ_{p,M}(l/K)\}$. That is, γ_i is the DFT of scaled β_l.
For $i = j$, property (4.14) follows from definition (4.11) of $Q_{p,M}$ and definition (4.13) of the Demmler-Reinsch basis. Hence, it holds that

$$\frac{1}{N}\sum_{l=1}^{N} |\phi_i(l/N)|^2 = \frac{N^{-1}\sum_{l=1}^N |\Phi_p\{Kl/N + (p+1)/2, \exp(-2\pi\mathtt{i}i/K)\}|^2}{Q_{p,M}(i/K)} = 1.$$

For $i \neq j$, the orthogonality of $\phi_i(x)$ and $\phi_j(x)$ follows from series representation (4.17) and the fact that $\sum_{l=1}^N \exp\{-2\pi\mathtt{i}l(i-j+sK)/N\} = 0$ and $i, j = 1, \ldots, K$.
To show property (4.15), one can again use representation (4.17) to find

$$\begin{aligned}\sqrt{Q_{p,M}(i/K)}\phi_i^{(q)}(x) &= (-2\pi\mathtt{i}i)^q \mathrm{sinc}(\pi i/K)^q \\ &\times \sum_{l=-\infty}^{\infty}(-1)^{lq}\mathrm{sinc}\{\pi(i/K+l)\}^{p+1-q}\exp\{-2\pi\mathtt{i}x(i+lK)\}.\end{aligned}$$

Since for $j, i = 1, \ldots, K$ and $\forall l, m \in \mathbb{Z}$ $\int_0^1 \exp\{-2\pi\mathtt{i}x(i-j+lK-mK)\}\,dx \neq 0$ iff $i = j$ and $l = m$, we obtain

$$Q_{p,M}(i/K)\int_0^1 \left|\phi_i^{(q)}(x)\right|^2 dx = (2\pi i)^{2q}\mathrm{sinc}(\pi i/K)^{2q}\sum_{l=-\infty}^{\infty}\mathrm{sinc}\left\{\pi\left(i/K+l\right)\right\}^{2p-2q+2},$$

which implies the assertion and proves the lemma. □

Remarks

1. Function $s \in S_{\mathtt{per}}(p; \underline{\tau}_K)$ iff $s(x) = \sum_{i=1}^K \phi_i(x)\beta_i$ with $\beta_i = \overline{\beta_{K-i}}$, $i = 1, \ldots, K/2$.

2. The continuous version of the Demmler-Reinsch basis and its eigenvalues is

$$\phi(u,x) := \Phi_p\{Kx + (p+1)/2, \exp(-2\pi\mathtt{i}u/K)\}/\sqrt{Q_{p,M}(u/K)} \quad (4.18)$$

$$\mu(u) := (2\pi u)^{2q}\mathrm{sinc}(\pi u/K)^{2q}Q_{2p-2q}(u/K)/Q_{p,M}(u/K). \quad (4.19)$$

For $u = 1, 2, \ldots K\, \phi(u, x) = \phi_u(x)$.

3. As discussed in the proof of Lemma 13, basis functions $\phi_i(x)$ have series representation (4.17) and $\phi_i(x)$ are the scaled DFT of periodic B-splines, i.e.

$$\phi_i(x) K \sqrt{Q_{p,M}(i/K)} = \sum_{l=1}^{K} B_{l,p}(x) \exp(-2\pi \mathtt{i} il/K).$$

4. Since $\phi_i(x)$ is the scaled DFT of a real-valued B-spline functions and $Q_{p,M}(z)$ is a symmetric, positive function by definition, it holds that $\phi_i(x) = \overline{\phi_{K-i}(x)}$ and $\mu_i = \mu_{K-i}$.

5. From series representation of Q-polynomials (4.7) and from series representation of the Demmler-Reinsch basis (4.17), it follows that

$$\int_0^1 \phi_i(x) \overline{\phi_j(x)} dx = \delta_{ij} Q_{2p}(i/K)/Q_{p,M}(i/K). \tag{4.20}$$

   In Lee and Tang (1991), Lee et al. (1992), Zheludev (1996), they essentially considered L_2-orthonormal basis $Q_{p,M}^{1/2}(i/K) Q_{2p}^{-1/2}(i/K) \phi_i(x), i = 1, \ldots, K$.

6. For $K = N$ and $p = 2q-1$, $\mu_i = (2\pi \mathtt{i})^{2q} \mathrm{sinc}(\pi i/K)^{2q} Q_{2q-2}(i/K)^{-1}$ and at the data points l/N, the Demmler-Reinsh basis reduces to $\phi_i(l/N) = \exp(-2\pi \mathtt{i} il)$.

7. Functions $\sqrt{2}\mathfrak{R}\{\phi_i(x)\}$, $\sqrt{2}\mathfrak{Im}\{\phi_i(x)\}$ for $i = 1, \ldots, K/2$ construct the real-valued Demmler-Reinsch basis.

Thus, any $s(x) \in S_{\mathtt{per}}(p; \underline{\tau}_K)$ can be represented as $s(x) = \sum_{i=1}^K \beta_i \phi_i(x)$ and the solution to (4.1) results in $\widehat{f}(x) = \sum_{i=1}^K \widehat{\beta}_i \phi_i(x)$ with $\widehat{\beta}_i = (1 + \lambda \mu_i)^{-1} \widehat{y}_i$, where

$$\widehat{y}_i \coloneqq N^{-1} \sum_{l=1}^{N} Y_l\, \overline{\phi_i(l/N)}. \tag{4.21}$$

From series representation (4.17) we obtain

$$\widehat{f}(x) = \frac{1}{\sqrt{Q_{p,M}(i/K)}} \sum_{l=-\infty}^{\infty} \sum_{i=1}^{K} \widehat{\beta}_i\, \mathrm{sinc}\{\pi(i/K + l)\}^{p+1} \exp\{-2\pi \mathtt{i} x(i + lK)\}.$$

Since $\widehat{\beta}_i = \widehat{\beta}_{i+lK}$ the Fourier coefficients of the spline estimator are given by

$$c_{i+lK} = \frac{\operatorname{sinc}\{\pi(i/K+l)\}^{p+1}}{\sqrt{Q_{p,M}(i/K)}}\widehat{\beta}_i = \frac{\operatorname{sinc}\{\pi(i/K+l)\}^{p+1}\widehat{y}_i}{\sqrt{Q_{p,M}(i/K)}(1+\lambda\mu_i)}, \tag{4.22}$$

where the c_l satisfy $\widehat{f}(x) = \sum_{l=-\infty}^{\infty} c_l \exp(-2\pi\mathtt{i}lx)$. From (4.22), one can immediately obtain the Fourier coefficients for both extreme cases: periodic smoothing and regression splines. In particular, for $\lambda = 0$, $K \ll N$ (periodic regression spline), it holds that

$$c_{i+lK} = \frac{\operatorname{sinc}\{\pi(i/K+l)\}^{p+1}\widehat{y}_i}{\sqrt{Q_{p,M}(i/K)}} = \frac{\operatorname{sinc}\{\pi(i/K+l)\}^{p+1}\widehat{y}_i}{\sqrt{Q_{2p}(i/K)+O(M^{-p-1})}}.$$

For $K = N$ and $p = 2q-1$ (periodic smoothing spline), it holds that

$$c_{i+lN} = \frac{\operatorname{sinc}\{\pi(i/N+l)\}^{2q}\check{y}_i}{Q_{2q-2}(i/N) + \lambda(2\pi i)^{2q}\operatorname{sinc}(\pi i/N)^{2q}},$$

with $\check{y}_i = N^{-1}\sum_{l=1}^{N}\exp(2\pi\mathtt{i}il/N)Y_l$. Note that for $\lambda = 0$, ratio $c_i/E\check{y}_i$ coincides with the i-th attenuation factor for spline interpolation obtained in Gautschi (1971) and discussed in Section 4.1.1.

4.2 The integrated mean squared error for periodic splines

In Section 2.5, we discussed the results of Claeskens et al. (2009) on the dependence of the global asymptotics (the AMSE) of penalized splines in the non-periodic case on parameter k_q. In this section, we obtain similar results, i.e two asymptotic scenarios depending on k_q, for the IMSE of periodic splines making use of the results obtained in the previous section. For periodic splines, we set

$$k_q := \lambda^{1/(2q)}\pi K. \tag{4.23}$$

Let us proceed with the study of the global asymptotics of periodic spine estimators. The IMSE of a spline estimator can be decomposed as

$$IMSE(\widehat{f}) = \int_0^1 \mathrm{Var}\{\widehat{f}(x)\}dx + \int_0^1 [E\{\widehat{f}(x)\} - s_{\mathtt{p}}(x)]^2 dx + \int_0^1 \{s_{\mathtt{p}}(x) - f(x)\}^2 dx,$$

where $s_{\mathtt{p}}$ is the best L_2-approximation of $f \in C^{p+1}_{\mathtt{per}}$ by $S_{\mathtt{per}}(p; \underline{\tau}_K)$. Thus, the IMSE consists of three summands: the integrated variance, the integrated squared shrinkage bias and the integrated squared approximation bias. The shrinkage bias appears due to the penalization involved in (4.1) and vanishes for regression splines ($\lambda = 0$). The approximation bias is the error caused by the approximation of a continuous function f by a spline. The sharp asymptotic behavior of the integrated squared approximation bias has been proved in Barrow and Smith (1978). In particular, they have shown that

$$\lim_{K\to\infty} K^{2p+2} \int_0^1 \{s_{\mathtt{p}}(x) - f(x)\}^2 dx = \frac{|\mathcal{B}_{2p+2}|}{(2p+2)!} \left(\int_0^1 |f^{p+1}(x)|^{1/(p+1.5)} dx \right)^{(2p+3)}. \quad (4.24)$$

The following theorem gives exact expressions for the integrated variance and the integrated squared shrinkage bias.

Theorem 1 *Let model (2.1) with assumptions (2.2) hold. Let* $f \in C^{p+1}_{\mathtt{per}}$ *and* $\widehat{f}(x) \in S_{\mathtt{per}}(p; \underline{\tau}_K)$ *be the solution to (4.1) with* $\{x_i = i/N\}_{i=1}^N$, *and* $\underline{\tau}_K = \{i/K\}_{i=0}^K$. *Then, the integrated squared shrinkage bias the integrated variance and the integrated squared shrinkage bias of* $\widehat{f}(x)$ *are given by*

$$\int_0^1 \mathrm{Var}\{\widehat{f}(x)\}dx = \frac{\sigma^2}{N} \sum_{i=1}^K \frac{Q_{2p}(i/K)}{Q_{p,M}(i/K)(1+\lambda\mu_i)^2} \quad (4.25)$$

$$\int_0^1 [\mathrm{E}\{\widehat{f}(x)\} - s_{\mathtt{p}}(x)]^2 dx = \sum_{i=1}^K \frac{Q_{2p}(i/K)(\lambda\mu_i)^2|f_i|^2}{Q_{p,M}(i/K)(1+\lambda\mu_i)^2} \left| 1 - \frac{f_{i,N} - f_i}{f_i \lambda \mu_i} \right|^2, \quad (4.26)$$

with $f_i = \sqrt{Q_{p,M}(i/K)/Q_{2p}(i/K)} \int_0^1 f(x)\phi_i(x)dx$ *and*
$f_{i,N} = \sqrt{Q_{p,M}(i/K)/Q_{2p}(i/K)} N^{-1} \sum_{l=1}^N f(l/N)\phi_i(l/N)$.

Proof of Theorem 1

From Parseval's identity, it holds that

$$
\begin{aligned}
\mathrm{Var}\{\widehat{f}(x)\} &= \sum_{i=1}^{K}\sum_{l=-\infty}^{\infty}\mathrm{Var}(c_{i+lK}) = \sum_{i=1}^{K}\sum_{l=-\infty}^{\infty}\frac{\mathrm{sinc}\{\pi(l+i/K)\}^{2p+2}}{Q_{p,M}(i/K)(1+\lambda\mu_i)^2}\mathrm{Var}(\widehat{y}_i) \\
&= \frac{\sigma^2}{N}\sum_{i=1}^{K}\frac{Q_{2p}(i/K)}{Q_{p,M}(i/K)(1+\lambda\mu_i)^2}.
\end{aligned}
$$

To obtain $Var(\widehat{y}_i) = N^{-1}\sigma^2$, we use (4.21) and $\sum_{l=1}^{N}|\phi_i(l/N)|^2 = N$. From (4.17) and (4.20), the projection estimator $s_{\mathtt{p}}(x)$ can be written as

$$
s_{\mathtt{p}}(x) = \sum_{i=1}^{K}\frac{\int_0^1 f(x)\phi_i(t)dt}{\int_0^1 \phi_i(t)\overline{\phi_i(t)}dt}\,\phi_i(x) = \sum_{i=1}^{K}\sum_{l=-\infty}^{\infty}\frac{\mathrm{sinc}\left\{\pi(i/K+l)\right\}^{p+1}f_i}{\sqrt{Q_{p,M}(i/K)}}\exp\left\{-2\pi\mathtt{i}x(i+lK)\right\},
$$

where

$$
f_i = Q_{p,M}^{1/2}(i/K)Q_{2p}^{-1/2}(i/K)\int_0^1 f(x)\phi_i(x)dx.
$$

With this,

$$
\begin{aligned}
&\int_0^1[\mathrm{E}\{\widehat{f}(x)\} - s_{\mathtt{p}}(x)]^2dx \\
&= \sum_{i=1}^{K}\sum_{l=-\infty}^{\infty}\left|E(c_{i+l/K}) - \mathrm{sinc}\{\pi(l+i/K)\}^{p+1}f_i/\sqrt{Q_{p,M}(i/K)}\right|^2 \\
&= \sum_{i=1}^{K}\sum_{l=-\infty}^{\infty}\frac{\mathrm{sinc}\{\pi(l+i/K)\}^{2p+2}}{Q_{p,M}(i/K)}\left|\frac{f_{i,N}}{1+\lambda\mu_i} - f_i\right|^2 \\
&= \sum_{i=1}^{K}\frac{Q_{2p}(i/K)(\lambda\mu_i)^2|f_i|^2}{Q_{p,M}(i/K)(1+\lambda\mu_i)^2}\left|1 - \frac{f_{i,N}/f_i - 1}{\lambda\mu_i}\right|^2,
\end{aligned}
$$

with $f_{i,N} = Q_{p,M}^{1/2}(i/K)Q_{2p}^{-1/2}(i/K)N^{-1}\sum_{l=1}^{N}f(l/N)\phi_i(l/N)$, proving the theorem. $\square$

Since the ratio $Q_{2p}(u)/Q_{p,M}(u)$ is bounded (see Section 4.1.2), from the equations (4.25) and (4.26), it is clear that the asymptotic behavior of spline-based estimators depends on

$$
\lambda\mu_{K/2} = k_q^{2q}(2/\pi)^{2q}Q_{2p-2q}(1/2)/Q_{p,M}(1/2),
$$

similar to the results in Claeskens et al. (2009). Since k_q differs only by a constant from $(\lambda\mu_{K/2})^{1/(2q)}$, the asymptotic behavior of spline-based estimators also depends on k_q. From (4.24), (4.25) and (4.26), one can find the asymptotic orders of the components of the IMSE in two asymptotic scenarios.

Corollary 1 *Under the assumptions of Theorem 1, for $p \geq 2q-1$, it holds that*

$$IMSE(\widehat{f}) = \begin{cases} O\left(KN^{-1}\right) + O\left(\lambda^2\right) + O(K^{-2p-2}), & \text{for } k_q = O(1), \\ O\left(\lambda^{-1/(2q)}N^{-1}\right) + O\left(\lambda^2\right) + O(K^{-2p-2}), & \text{for } k_q \to \infty, \end{cases}$$

so that for $k_q = O(1)$ and $K \asymp N^{1/(2p+3)}$, $\lambda = O(N^{-\nu})$, $\nu \in [(p+1)/(2p+3), 1]$ imply $IMSE(\widehat{f}) = O\left\{N^{-(2p+2)/(2p+3)}\right\}$. For $k_q \to \infty$ and $\lambda = O\left\{N^{-2q/(4q+1)}\right\}$ with $\lambda N \to \infty$, $K \asymp N^{\varsigma}$, $\varsigma \in [1/(4q+1), 1]$ it holds that $IMSE(\widehat{f}) = O\left\{N^{-4q/(4q+1)}\right\}$.

Proof of Corollary 1
Let $k_q = O(1)$. Then, since $Q_{2p}(i/K)/Q_{p,M}(i/K) \leq 3/2$ for all i,

$$\begin{aligned} \mathrm{Var}\{\widehat{f}(x)\} &= 2\frac{\sigma^2}{N}\sum_{i=1}^{K/2}\frac{Q_{2p}(i/K)}{Q_{p,M}(i/K)(1+\lambda\mu_i)^2} < 3\frac{\sigma^2}{N}\sum_{i=1}^{K/2}\frac{1}{(1+\lambda\mu_i)^2} \\ &< 3/2\frac{\sigma^2}{N}K = O(KN^{-1}). \end{aligned}$$

In case $k_q \to \infty$, we use the bound $\mu_i \geq (2\pi i)^{2q},\ i = 1, \ldots, K/2$ that follows from (4.16) and

$$\mathrm{sinc}(\pi x)^{2q}\frac{Q_{2p-2q}(x)}{Q_{p,M}(x)} \geq \mathrm{sinc}(0)^{2q}\frac{Q_{2p-2q}(0)}{Q_{p,M}(0)} = 1,\ x \in [0, 1/2].$$

Using this bound and approximating sums by integrals as in Wahba (1975), we bound the integrated variance by

$$\begin{aligned} \mathrm{Var}\{\widehat{f}(x)\} &\leq \frac{3\sigma^2}{N}\sum_{i=1}^{K/2}\frac{1}{\{1+\lambda(2\pi i)^{2q}\}^2} = \frac{3\sigma^2}{N}\left[\int_0^{K/2}\frac{dx}{\{1+\lambda(2\pi x)^{2q}\}^2} - 1 + r_q\right] \\ &= \frac{3\sigma^2}{N}\left\{\int_0^{k_q}\frac{\lambda^{-1/(2q)}dx}{2\pi\left(1+x^{2q}\right)^2} - 1 + r_q\right\} \leq \frac{3\sigma^2}{N}\left\{\int_0^{\infty}\frac{\lambda^{-1/(2q)}dx}{2\pi\left(1+x^{2q}\right)^2} - 1 + r_q\right\}, \end{aligned}$$

where $r_q = O(1)$ is the remainder term of the Euler-Maclaurin formula, $\int_0^\infty \left(1+x^{2q}\right)^{-2} dx$ is a positive constant that can be found for each q. Hence, if $\lambda^{1/(2q)}N \to \infty$, the rate of the integrated variance is $O\left\{\lambda^{-1/(2q)}N^{-1}\right\}$.

To find the rate of the integrated squared shrinkage bias, we will need the quadrature errors between $f_{i,N}$ and f_i. From Utreras (1980), $\forall g_1, g_2 \in W^q[0,1]$ $\exists c > 0$ such that

$$\begin{aligned}&\left|\frac{1}{N}\sum_{i=1}^{N} g_1(i/N)g_2(i/N) - \int_0^1 g_1(x)g_2(x)dx\right|\\ \leq& N^{-2}c\int_0^1 \left\{g_1^{(q)}(x)\right\}^2 dx \int_0^1 \left\{g_2^{(q)}(x)\right\}^2 dx.\end{aligned}$$

Using this quadrature error and the fact that $f, \mathfrak{R}\left(\phi_i\right), \mathfrak{Im}\left(\phi_i\right) \in W^q[0,1]$, we obtain

$$\begin{aligned}\left|f_{i,N} - f_i\right| \leq& \frac{Q_{p,M}^{1/2}(i/K)}{Q_{2p}(i/K)}\left|\frac{1}{N}\sum_{l=1}^{N} f(l/N)\mathfrak{R}\left\{\phi_i(l/N)\right\} - \int_0^1 f(x)\mathfrak{R}\left\{\phi_i(x)\right\}dx\right|\\ &+\frac{Q_{p,M}^{1/2}(i/K)}{Q_{2p}(i/K)}\left|\frac{1}{N}\sum_{l=1}^{N} f(l/N)\mathfrak{Im}\left\{\phi_i(l/N)\right\} - \int_0^1 f(x)\mathfrak{Im}\left\{\phi_i(x)\right\}dx\right|\\ &=O\left(N^{-2}\right).\end{aligned}$$

Similarly, $\left|f_{i,N}/f_i - 1\right| = O(N^{-1})$.

Further, let us consider the integrated squared shrinkage bias in different scenarios. If $k_q = O(1)$, we bound the integrated shrinkage bias (4.26) as follows.

$$\begin{aligned}\int_0^1 [\mathrm{E}\{\widehat{f}(x)\} - s_{\mathtt{p}}(x)]^2 dx =& \sum_{i=1}^{K}\frac{Q_{2p}(i/K)}{Q_{p,M}(i/K)}\left|\frac{f_{i,N}}{1+\lambda\mu_i} - f_i\right|^2 \leq 3\sum_{i=1}^{K/2}\left|f_{i,N} - f_i - \frac{\lambda\mu_i f_{i,N}}{1+\lambda\mu_i}\right|^2\\ \leq& 3\sum_{i=1}^{K/2}\left|\left|f_{i,N} - f_i\right| + \lambda\mu_i\left|f_{i,N}\right|\right|^2 = O(K/N) + O(\lambda^2).\end{aligned}$$

If $k_q \to \infty$, the integrated squared shrinkage bias (4.26) can be bounded by

$$\begin{aligned}
\int_0^1 [\mathrm{E}\{\widehat{f}(x)\} - s_{\mathrm{p}}(x)]^2 dx &= \sum_{i=1}^{K} \frac{Q_{2p}(i/K)(\lambda\mu_i)^2|f_i|^2}{Q_{p,M}(i/K)(1+\lambda\mu_i)^2} \left|1 - \frac{f_{i,N}/f_i - 1}{\lambda\mu_i}\right|^2 \\
&\leq 3\lambda^2 \sum_{i=1}^{K/2} \mu_i^2 |f_i|^2 \left|1 - \frac{f_{i,N}/f_i - 1}{\lambda\mu_i}\right|^2 \\
&\leq 3\lambda^2 \sum_{i=1}^{K/2} \mu_i^2 |f_i|^2 \{1 + O(N^{-1})\} = O(\lambda^2).
\end{aligned}$$

Optimizing the $IMSE(\widehat{f})$ with respect to the parameters K and λ gives the optimal rates for K and λ in both asymptotic scenarios. □

Corollary 1 states that depending on k_q, and thus on the number of knots K taken, the asymptotic scenario of periodic spline estimators is either similar to the periodic regression spline asymptotics or to the periodic smoothing spline asymptotics. For small $K \asymp N^{1/(2p+3)}$ with $k_q = O(1)$, the convergence rate of the estimator $N^{-(p+1)/(2p+3)}$ is the same as that for the regression splines (see Zhou et al., 1998) and λ is, in fact, non-identifiable, i.e. can not be estimated consistently. Once more knots are taken, so that $k_q \to \infty$, the smoothing parameter λ controls the fit and the convergence rate is $N^{-2q/(4q+1)}$, as was found for periodic smoothing spline estimators by Wahba (1975). In this scenario, K is non-identifiable, meaning that taking any K satisfying $K \asymp N^{\varsigma}$, $\varsigma \in [1/(4q+1), 1]$ has no influence on the IMSE. Apparently, the choice of p and q is important for the convergence rate in each scenario. Taking $p > 2q - 1$ leads to a faster convergence rate in the "small" number of knots scenario, while for $p = 2q - 1$ the convergence rate in both scenarios is the same.

4.3 Equivalent kernel

Periodic splines are linear in observations and can be written as a weighted sum

$$\widehat{f}(x) = \frac{1}{N} \sum_{i=1}^{N} W(x, i/N) Y_i.$$

We refer to function $W(x,t)$ as the equivalent kernel for periodic spline estimators. Strictly speaking, $W(x,t)$ is an effective kernel, since it depends on observation points $\{i/N\}_{i=1}^N$. However, in case of periodic splines, equidistant knots and observations, this dependence is simpler than in general case and can be expressed via parameter $M = N/K$ – the number of observations between two subsequent knots. Thus, in contrast to the available results on smoothing and regression splines, we work with the exact form of $W(x,t)$ avoiding any asymptotic approximation of the effective kernel (and herewith of the estimator).
Knowing the properties and the exact form of the equivalent kernel allows us to study the pointwise asymptotic behaviour of periodic splines. In this section, we study the periodic equivalent kernel on the interval and a related non-periodic kernel on the real line.

4.3.1 Periodic equivalent kernels and equivalent kernels on real line

Using the Demmler-Reinsch basis, we can write the solution to (4.1) as

$$\widehat{f}(x) = \frac{1}{N}\sum_{l=1}^{N}\sum_{i=1}^{K}\frac{\phi_i(x)\overline{\phi_i(l/N)}}{1+\lambda\mu_i}Y_l. \tag{4.27}$$

Hence, the equivalent kernel for a periodic spline estimator is

$$W(x,t) = \sum_{i=1}^{K}\frac{\overline{\phi_i(t)}}{1+\lambda\mu_i}\phi_i(x). \tag{4.28}$$

Some properties of the equivalent kernel follow immediately.

1. $W(x,t)$ is a real-valued function $\forall x,t\in[0,1]$.
2. $W(x,t)\in S_{\mathtt{per}}(p;\mathcal{T}_K)$ at a fixed t (or fixed x).
3. $W(x,t)$ is the reproducing kernel of $S_{\mathtt{per}}(p;\mathcal{T}_K)$ with respect to inner product

$$\langle f,g\rangle_\lambda \coloneqq N^{-1}\sum_{i=1}^{N} f(i/N)g(i/N) + \lambda\int_0^1 f^{(q)}(x)g^{(q)}(x)dx,$$

i.e. $\langle W(t,x),s(x)\rangle_\lambda = s(t),\ \forall s\in S_{\mathtt{per}}(p;\mathcal{T}_K)$ and $t\in[0,1]$ for any $N\geq K$.

Properties 1, 2 follow from the fact, that the i-th and the $(K-i)$-th summands in (4.28) are complex conjugate. Property 3 follows from Lemma 13 and the representation of splines in the Demmler-Reinsch basis.
We define a kernel on the real line via

$$\mathcal{W}(x,t) := \int_0^K \frac{\phi(u,x)\overline{\phi(u,t)}}{1+\lambda\mu(u)}du, \tag{4.29}$$

where function $\phi(u,x)$ and $\mu(u)$ are given in (4.18) and (4.19) correspondingly. Note, $\mathcal{W}(x,t)$ is a real valued function, since $\phi(u,x) = \overline{\phi(K-u,x)}$ and $\mu(u) = \overline{\mu(K-u)}$.

Lemma 14 *Equivalent kernel $W(x,t)$ can be obtained by "folding back" $\mathcal{W}(x,t)$, that is*

$$W(x,t) = \sum_{l=-\infty}^{\infty} \mathcal{W}(x,t+l). \tag{4.30}$$

Proof of Lemma 14
Note that for fixed x, function $\sum_{l=-\infty}^{\infty} \mathcal{W}(x,t+l)$ is a periodic real-valued function of t with period 1. Hence, it is enough to show that for fixed x, the Fourier coefficients of $W(x,t)$ and $\sum_{l=-\infty}^{\infty} \mathcal{W}(x,t+l)$ coincide. From now on, we consider both functions as functions of t, while x is being a fixed parameter.
Substituting series representation (4.17) for $\overline{\phi_i(t)}$ into (4.28), we obtain the Fourier series of $W(x,t)$:

$$W(x,t) = \sum_{l=-\infty}^{\infty}\sum_{i=1}^{K} \frac{\text{sinc}\{\pi(i/K+l)\}^{p+1}\phi_i(x)}{\sqrt{Q_{p,M}(i/K)}(1+\lambda\mu_i)} \exp\{2\pi \mathtt{i} t(i+lK)\}.$$

Next, we use $Q_{p,M}(i/K) = Q_{p,M}(i/K+l)$, $\mu_i = \mu_{i+lK}$ and $\phi_i(x) = \phi_{i+lK}(x)$ for $l \in \mathbb{Z}$. Similar to the Fourier coefficients c_{i+lK} of $\widehat{f}(x)$ as given in Section 4.1.3, the Fourier coefficients $a_l(x)$ of $W(x,t)$ are given by

$$a_l(x) = \frac{\text{sinc}\{\pi(l/K)\}^{p+1}\phi_l(x)}{\sqrt{Q_{p,M}(l/K)}(1+\lambda\mu_l)},\ l \in \mathbb{Z}, \tag{4.31}$$

for $W(x,t) = \sum_{l=-\infty}^{\infty} a_l(x)\exp(2\pi \mathtt{i} lt)$. From the Poisson summation formula

$$\int_0^1 \sum_{j=-\infty}^{\infty} \mathcal{W}(x, t+j)\exp(-2\pi \mathtt{i} tl)dt = \int_{-\infty}^{\infty} \mathcal{W}(x,t)\exp(-2\pi itl)dt, \tag{4.32}$$

follows the equality of l-th Fourier coefficients of $\sum_{j=-\infty}^{\infty} \mathcal{W}(x, t+j)$ and of the Fourier transform of $\mathcal{W}(x,t)$. Applying the Poisson summation formula again, we obtain

$$\begin{aligned} \mathcal{W}(x,t) &= \int_0^K \sum_{l=-\infty}^{\infty} \frac{\mathrm{sinc}\{\pi(u/K+l)\}^{p+1}\phi(u,x)}{\sqrt{Q_{p,M}(u/K)}\{1+\lambda\mu(u)\}} \exp\{2\pi \mathtt{i} t(u+lK)\}du \\ &= \int_{-\infty}^{\infty} \frac{\mathrm{sinc}\{\pi(u/K)\}^{p+1}\phi(u,x)}{\sqrt{Q_{p,M}(u/K)}\{1+\lambda\mu(u)\}} \exp(2\pi \mathtt{i} tu)du. \end{aligned} \tag{4.33}$$

Using the inverse Fourier transform for (4.33), we obtain

$$\int_{-\infty}^{\infty} \mathcal{W}(x,t)\exp(-2\pi \mathtt{i} tu)dt = \frac{\mathrm{sinc}\{\pi(u/K)\}^{p+1}\phi(u,x)}{\sqrt{Q_{p,M}(u/K)}\{1+\lambda\mu(u)\}}. \tag{4.34}$$

If we take in (4.34) $u = l \in \mathbb{Z}$, then from (4.32) follows the equality of the l-th Fourier coefficients of $\sum_{j=-\infty}^{\infty} \mathcal{W}(x,t+j)$ and $a_l(x)$ in (4.31), which proves (4.30). □

4.3.2 Explicit formula for the equivalent kernel

In this section, we give the explicit form of equivalent kernels $W(x,t)$ and $\mathcal{W}(x,t)$. The obtained solutions depend on the roots of a certain combination of Euler-Frobenius polynomials given in Theorem 2. The equivalent kernels look complicated, but they are exact and hold for any p, q, M, K and λ. Related results in term of B-splines are to be found in Huang and Studden (1990) for regression splines and in Li and Ruppert (2008) for penalized splines when $p = q = 1$. Both worked with an approximation of the weight function of spline estimators.

Theorem 2 *Let*

$$P_{2p}(u) := \widetilde{\Pi}_{p,M}(u) + (-1)^q (k_q \pi^{-1})^{2q}(1-u)^{2q}\Pi_{2p-2q+1}(u)/(2p-2q+1)! \tag{4.35}$$

be a polynomial of degree $2p$ where $\Pi_{2p-2q+1}(u)$ is the Euler-Frobenius polynomial and

$\widetilde{\Pi}_{p,M}(u)$ be a linear combination of the Euler-Frobenius polynomials given in Lemma 12. Let also r_j, r_j^{-1}, $j = 1, \ldots, p$ be the roots of $P_{2p}(u)$ with $|r_j| < 1$. Then, denoting $P'_{2p}(r_j) \coloneqq \partial P_{2p}(u)/\partial u|_{u=r_j}$, $d_{x,t} \coloneqq \lfloor Kx - \{\frac{p+1}{2}\}\rfloor - \lfloor Kt - \{\frac{p+1}{2}\}\rfloor$ and representing

$$z^{p-d_{x,t}}\Phi_p\left\{Kx + (p+1)/2, z\right\}\Phi_p\left\{Kt + (p+1)/2, z^{-1}\right\} = \sum_{l=0}^{2p} \alpha_l\left(\{Kx\}, \{Kt\}\right) z^l$$

for some functions $\alpha_l(t_1, t_2)$ and $x, t \in \mathbb{R}$, we find

$$\begin{aligned} W(x,t) &= K\sum_{j=1}^{p}\sum_{l=0}^{2p}\frac{\alpha_l\left(\{Kx\},\{Kt\}\right)}{P'_{2p}(r_j)}\,\frac{r_j^{(d_{x,t}+l-1) \bmod K} + r_j^{K+2p-2-(d_{x,t}+l-1) \bmod K}}{\left(1-r_j^K\right)} \\ \mathcal{W}(x,t) &= K\sum_{j=1}^{p}\sum_{l=0}^{2p}\frac{\alpha_l\left(\{Kx\},\{Kt\}\right)}{P'_{2p}(r_j)}\,r_j^{|d_{x,t}+l-1|+I_{\{d_{x,t}\le -l\}}(2p-2)}. \end{aligned} \tag{4.36}$$

Proof of Theorem 2

Using formulas of Demmler-Reinsch basis (4.13), its eigenvalues (4.16) and formulas (4.23), (4.28), (4.29) we can represent functions $W(x,t)$ and $\mathcal{W}(x,t)$ as

$$\begin{aligned} W(x,t) &= \sum_{i=1}^{K}\frac{\Phi_p\left\{Kx+(p+1)/2,\exp(-2\pi\mathtt{i}i/K)\right\}\Phi_p\left\{Kt+(p+1)/2,\exp(2\pi\mathtt{i}i/K)\right\}}{Q_{p,M}(i/K)+\left(2k_q\pi^{-1}\right)^{2q}\sin(\pi i/K)^{2q}Q_{2p-2q}(i/K)} \\ \mathcal{W}(x,t) &= \int_0^K \frac{\Phi_p\left\{Kx+(p+1)/2,\exp(-2\pi\mathtt{i}u/K)\right\}\Phi_p\left\{Kt+(p+1)/2,\exp(2\pi\mathtt{i}u/K)\right\}}{Q_{p,M}(u/K)+\left(2k_q\pi^{-1}\right)^{2q}\sin(\pi u/K)^{2q}Q_{2p-2q}(u/K)}du. \end{aligned}$$

We aim to represent $W(x,t)$ and $\mathcal{W}(x,t)$ as a ratio of two polynomials in exponential function. From (4.9), it holds that

$$Q_{2p-2q}(u) = \exp\left\{2\pi\mathtt{i}u(p-q)\right\}\Pi_{2p-2q+1}\left\{\exp(-2\pi\mathtt{i}u)\right\}/(2p-2q+1)!.$$

Moreover, from Lemma 12 we know that $Q_{p,M}(u) = \exp(2\pi\mathtt{i}pu)\widetilde{\Pi}_{p,M}\{\exp(-2\pi\mathtt{i}u)\}$. With this, the denominators of $W(x,t)$ and $\mathcal{W}(x,t)$ can be expressed in terms of exponential function using

$$Q_{p,M}(u) + (2k_q\pi^{-1})^{2q}\sin(\pi u)^{2q}Q_{2p-2q}(u) = \exp(2\pi\mathtt{i}pu)P_{2p}\left\{\exp(-2\pi\mathtt{i}u)\right\}. \tag{4.37}$$

Using the representation of the exponential splines in terms of polynomials $\alpha_l(t_1, t_2)$ given in the theorem, we can express $W(x,t)$ and $\mathcal{W}(x,t)$ as a ratio of two polynomials in exponential functions:

$$\begin{aligned} W(x,t) &= \sum_{i=1}^{K} \frac{\exp\left(-2\pi \mathtt{i} d_{x,t} i/K\right) \sum_{l=0}^{2p} \alpha_l(\{Kx\},\{Kt\}) \exp\left(-2\pi \mathtt{i} il/K\right)}{P_{2p}\left\{\exp\left(-2\pi \mathtt{i} i/K\right)\right\}} \\ \mathcal{W}(x,t) &= \int_0^K \frac{\exp\left(-2\pi \mathtt{i} d_{x,t} u/K\right) \sum_{l=0}^{2p} \alpha_l(\{Kx\},\{Kt\}) \exp\left(-2\pi \mathtt{i} ul/K\right)}{P_{2p}\left\{\exp\left(-2\pi \mathtt{i} u/K\right)\right\}} du. \end{aligned}$$

Next, we find the partial fractional decomposition of $1/P_{2p}(u)$ that allows us to use the inverse discrete Fourier transform in order to find $W(x,t)$ and methods of contour integration in order to find $\mathcal{W}(x,t)$. Before we proceed, we need to study roots of polynomial $P_{2p}(u)$. Define additionally

$$\widetilde{P}_p\{\cos(\pi u)^2\} \coloneqq \exp(2\pi \mathtt{i} pu) P_{2p}\left\{\exp(-2\pi \mathtt{i} u)\right\}. \tag{4.38}$$

Function $\widetilde{P}_p(u)$ is a polynomial of degree p whose existence follows from (4.37) and the fact that functions $Q_{2p-2q}(u)$ and $Q_{p,M}(u)$ can be expressed as polynomials of $\cos(\pi u)^2$ (see Section 4.1.1 and 4.1.2). Equation (4.38) is equivalent to the representation

$$P_{2p}(u) = u^p \widetilde{P}_p\left(\frac{u + u^{-1} + 2}{4}\right), \tag{4.39}$$

since $\cos(\pi u)^2 = \{\exp\left(-2\pi \mathtt{i} u\right) + \exp\left(2\pi \mathtt{i} u\right) + 2\}/4$. Hence, the roots of $P_{2p}(u)$ can be found from the roots of polynomial a smaller degree $\widetilde{P}_p(u)$. Moreover, the inverse of a root of $P_{2p}(u)$ is also a root of $P_{2p}(u)$, i.e. the roots of $P_{2p}(u)$ can be written as $r_j, r_j^{-1},\ i = 1, \ldots, p$ with $|r_j| < 1$. Note that roots of the Euler-Frobenius polynomial have the same property (see Section 4.1.1).

Now, the coefficients of the partial fractional decomposition of $1/P_{2p}(u)$ are $1/P_{2p}^{'}(r_j)$ and $1/P_{2p}^{'}\left(r_j^{-1}\right)$ correspond to roots r_j and r_j^{-1} for $j = 1, \ldots, p$. Further, from (4.39), we obtain

$$P_{2p}'(u) = pu^{p+1}\tilde{P}_p\left(\frac{u + u^{-1} + 2}{4}\right) + u^{p-1}\frac{u - u^{-1}}{4}\tilde{P}_p'\left(\frac{u + u^{-1} + 2}{4}\right)$$

and therefore

$$\begin{aligned} P'_{2p}(r_i) =& r_i^{p-1}\frac{r_i - r_i^{-1}}{4}\tilde{P}'_p\left(\frac{r_i + r_i^{-1}+2}{4}\right) \\ P'_{2p}(r_i^{-1}) =& -r_i^{1-p}\frac{r_i - r_i^{-1}}{4}\tilde{P}'_p\left(\frac{r_i + r_i^{-1}+2}{4}\right). \end{aligned}$$

Hence, making use of $P'_{2p}(r_i^{-1}) = -r_i^{2-2p}P'_{2p}(r_i)$, we obtain

$$\begin{aligned} W(x,t) =& \sum_{j=1}^{p}\sum_{l=0}^{2p}\frac{\alpha_l\left(\{Kx\},\{Kt\}\right)}{P'_{2p}\left(r_j\right)}R_n(j,l) \\ \mathcal{W}(x,t) =& \sum_{j=1}^{p}\sum_{l=0}^{2p}\frac{\alpha_l(\{Kx\},\{Kt\})}{P'_{2p}\left(r_j\right)}R(j,l), \end{aligned}$$

for

$$\begin{aligned} R_n(j,l) &\coloneqq \sum_{i=1}^{K}\left[\frac{\exp\left\{-2\pi\mathtt{i}\left(d_{x,t}+l\right)i/K\right\}}{\exp(-2\pi\mathtt{i}i/K)-r_j} - \frac{r_j^{2p-2}\exp\left\{-2\pi\mathtt{i}\left(d_{x,t}+l\right)i/K\right\}}{\exp(-2\pi\mathtt{i}i/K)-r_j^{-1}}\right] \\ R(j,l) &\coloneqq \int_0^K\left[\frac{\exp\left\{-2\pi\mathtt{i}\left(d_{x,t}+l\right)u/K\right\}}{\exp(-2\pi\mathtt{i}u/K)-r_j} - \frac{r_j^{2p-2}\exp\left\{-2\pi\mathtt{i}\left(d_{x,t}+l\right)u/K\right\}}{\exp(-2\pi\mathtt{i}u/K)-r_j^{-1}}\right]du. \end{aligned}$$

To find $R_n(j,l)$, we use the inverse discrete Fourier transform of the sequence $\{r_j^i\}_{i=1}^K$, which can be obtained from the geometric progression formula, so that

$$R_n(j,l) = K\ \frac{r_j^{(d_{x,t}+l-1)\bmod K} + r_j^{K+2p-2-(d_{x,t}+l-1)\bmod K}}{1-r_j^K}.$$

To find $R(j,l)$, we use of a change of variable $\exp(-2\pi iu/K) = z$ for $d_{x,t}+l > 0$ and $\exp(2\pi iu/K) = z$ for $d_{x,t}+l \leq 0$. Then, we apply the Cauchy integral formula

$$R(j,l) = \begin{cases} \frac{K}{2\pi\mathtt{i}}\oint_{|z|=1}\left(\frac{z^{d_{x,t}+l-1}}{z-r_j} - \frac{r_j^{2p-2}z^{d_{x,t}+l-1}}{z-r_j^{-1}}\right)dz = Kr_j^{d_{x,t}+l-1}, & \text{for } d_{x,t}+l>0 \\ \frac{K}{2\pi\mathtt{i}}\oint_{|z|=1}\left(\frac{r_j^{2p-1}z^{-d_{x,t}-l}}{z-r_j} - \frac{r_j^{-1}z^{-d_{x,t}-l}}{z-r_j^{-1}}\right)dz = Kr_j^{-d_{x,t}-l+2p-1}, & \text{for } d_{x,t}+l\leq 0, \end{cases}$$

where the contour integral is taken counter-clockwise. □

There is a closed-form expression available for the roots of polynomial P_{2p} for degrees $p \leq 4$. For these degrees, both $W(x,t)$ and $\mathcal{W}(x,t)$ can be obtained explicitly.
For $p = q = 1$, both functions have simple representations that become increasingly involved for larger p and q. For $p = q = 1$, we find

$$r_1 = 1 - \left(\sqrt{6/M^2 + 3 + 36k_q^2\pi^{-2}} - 3\right)/(6k_q^2\pi^{-2} - 1 + 1/M^2),$$

$P_2'(r_1) = \sqrt{(2/M^2 + 1 + 12k_q^2\pi^{-2})/3}$, $\alpha_2(t_1,t_2) = \alpha_0(t_2,t_1) = t_1 - t_1t_2$ and $\alpha_1(t_1,t_2) = 1 - \alpha_0(t_1,t_2) - \alpha_2(t_1,t_2)$. Hence, for $p = 1$ the formula for $\mathcal{W}(x,t)$ is

$$\begin{aligned} \mathcal{W}(x,t) &= K\sqrt{3/\left(2M^{-2} + 1 + 12k_q^2\pi^{-2}\right)}\Big\{(\{Kt\} - \{Kx\}\{Kt\})\, r_1^{|\lfloor Kx\rfloor - \lfloor Kt\rfloor - 1|} \\ &+ (1 - \{Kx\} - \{Kt\} + 2\{Kx\}\{Kt\})\, r_1^{|\lfloor Kx\rfloor - \lfloor Kt\rfloor|} \\ &+ (\{Kx\} - \{Kx\}\{Kt\})\, r_1^{|\lfloor Kx\rfloor - \lfloor Kt\rfloor + 1|}\Big\}. \end{aligned}$$

For $k_q = 0$ and $M \to \infty$, the root is $r_1 = -2 + \sqrt{3}$ and kernel $\mathcal{W}(x,t)$, scaled with K, coincides with the equivalent kernel for regression splines $\mathcal{K}_{rs}(x,t)$, obtained in Huang and Studden (1990):

$$\begin{aligned} \mathcal{K}_{rs}(x,t) &= \sqrt{3}\Big\{(\{t\} - \{x\}\{t\})\, r_1^{|\lfloor x\rfloor - \lfloor t\rfloor - 1|} + (1 - \{x\} - \{t\} + 2\{x\}\{t\})\, r_1^{|\lfloor x\rfloor - \lfloor t\rfloor|} \\ &+ (\{x\} - \{x\}\{t\})\, r_1^{|\lfloor x\rfloor - \lfloor t\rfloor + 1|}\Big\}. \end{aligned}$$

The last formula can be found by plugging (2.3) into (3.26) for $p = 1$ and some simplification. For $k_q \to \infty$, $M = 1$, we scale $\mathcal{W}(x,t)$ with $\pi\lambda^{1/(2q)}$ and make use of $\lim_{k_q\to\infty} r_1^{k_q|x-t|} = \exp(-\pi\,|x-t|)$, $\lim_{k_q\to\infty} r_1 = 1$ obtaining for $p = 1$ that

$$\lim_{k_q\to\infty}\left[\lambda^{1/(2q)}\mathcal{W}\left\{\lambda^{1/(2q)}x, \lambda^{1/(2q)}t\right\}\right] = 1/2\exp(-|x-t|) = \mathcal{K}_{ss}(x-t).$$

For $p = 3$, $q = 2$, the explicit form of the equivalent kernel is a lot more complicated. The $\alpha_l(t_1, t_2)$ coefficients are given by

$$\begin{aligned}
\alpha_0(t_1,t_2) &= (t_2 - t_1 t_2)^3/36\\
\alpha_1(t_1,t_2) &= (1-t_1)^3(1+3t_2+3t_2^2)/36 + (1+9t_1-15t_1^2+6t_1^3)t_2^3/36\\
\alpha_2(t_1,t_2) &= \frac{(2-t_1)^3}{36} + \frac{4-6t_1^2+3t_1^3}{12t_2^{-1}} - \frac{8+6t_1-30t_1^2+15t_1^3}{36t_2^{-3}} + \frac{2+6t_1-12t_1^2+5t_1^3}{12t_2^{-2}}\\
\alpha_3(t_1,t_2) &= \frac{1}{2} + \frac{t_1^2(4t_1-9)}{18} + \frac{t_1(3-t_1^2)t_2 - (3-9t_1^2+5t_1^3)t_2^2}{6} - \frac{(2t_1-1)(4+5t_1-5t_1^2)t_2^3}{18}
\end{aligned}$$

and $\alpha_6(t_1,t_2) = \alpha_0(t_2,t_1)$, $\alpha_5(t_1,t_2) = \alpha_1(t_2,t_1)$, $\alpha_4(t_1,t_2) = \alpha_2(t_2,t_1)$. Additionally, let

$$\begin{aligned}
a &\coloneqq 4/315 - 8/189M^{-6} + 4/135M^{-4} + 32/3k_q^4\pi^{-4}\\
b &\coloneqq 38/105 + 8/63M^{-6} - 2/45M^{-4} - 16k_q^4\pi^{-4}\\
c &\coloneqq 4/7 - 8/63M^{-6}\\
d &\coloneqq 17/315 + 8/189M^{-6} + 2/135M^{-4} + 16/3k_q^4\pi^{-4}\\
V &\coloneqq 2^{-1/3}\left\{\sqrt{(2b^3-9abc+27a^2d)^2 - 4(b^2-3ac)^3} + 2b^3 - 9abc + 27a^2d\right\}^{1/3}.
\end{aligned}$$

Then the root of cubic polynomial $\widetilde{P}_p(u)$ used in the proof of the theorem are given by

$$\begin{aligned}
R_1 &\coloneqq -\left\{b + V + \left(b^2-3ac\right)V^{-1}\right\}/(3a)\\
R_2 &\coloneqq \left\{-b + \left(1+\mathtt{i}\sqrt{3}\right)V/2 + \left(1-\mathtt{i}\sqrt{3}\right)\left(b^2-3ac\right)V^{-1}/2\right\}/(3a)\\
R_3 &\coloneqq \left\{-b + \left(1-\mathtt{i}\sqrt{3}\right)V/2 + \left(1+\mathtt{i}\sqrt{3}\right)\left(b^2-3ac\right)V^{-1}/2\right\}/(3a).
\end{aligned}$$

The formula for the roots r_i and $P_{2p}'(r_j)$ for $i = 1, 2, 3$ is

$$\begin{aligned}
r_i &= \min\left\{-1+2R_i+2\sqrt{R_i^2-R_i},\ \left(-1+2R_i+2\sqrt{R_i^2-R_i}\right)^{-1}\right\}\\
P_{2p}'(r_i) &= \left(r_i^3 - r_i\right)\left(3aR_i^2 + 2bR_i + c\right)/4.
\end{aligned}$$

Note that root $\{r_i\}_{i=1}^3$ depend on k_q and M. Even though the formula if complicated, the exact values of $\{r_i\}_{i=1}^3$ and $\left\{P_{2p}'(r_i)\right\}_{i=1}^3$ can be computed for any fixed k_q and M. In particular, for $k_q = 0$ and $M \to \infty$, we obtain $r_1 \simeq -0.5353$, $r_2 \simeq -0.1226$, $r_3 \simeq -0.0091$, $P'_{2p}(r_1) \simeq 0.0476$, $P'_{2p}(r_2) \simeq -0.0142$, $P'_{2p}(r_3) \simeq 0.0196$.

4.4 Local asymptotic properties of spline estimators

In this section, we find the precise pointwise behavior of periodic spline estimators. For this, we need to introduce the scaled equivalent kernel on $\mathbb{R}$, defined as

$$\mathcal{K}(x,t) := h(k_q)\mathcal{W}\left\{h(k_q)x, h(k_q)t\right\},$$

where $h(k_q)$ is a bandwidth given in Section 4.4.1.

4.4.1 Bandwidth

We introduce a bandwidth, which is universal for all periodic spline estimators. Let us motivate the choice of the bandwidth. We want the bandwidth to satisfy the following conditions

1. It is a smooth function of k_q.
2. It tends to zero as the number of observation grows.
3. For $k_q \to \infty$, it coincides up to a constant with bandwidth $\lambda^{1/(2q)}$ of the equivalent kernel for smoothing splines.
4. For $k_q = 0$, it coincides with the bandwidth $1/K$ of the equivalent kernel for regression splines.

Let us define the bandwidth and check its properties.

$$h(k_q)^{-1} := \int_0^K \frac{dx}{1+\lambda(\pi x)^{2q}} = \lambda^{-1/(2q)}\pi^{-1}\int_0^{k_q} \frac{dx}{1+x^{2q}}.$$

Bandwidth $h(k_q)$ is a smooth function of k_q with a (rather complicated) closed form expression available for each q. However, for our subsequent developments the following representation will be more suitable.

$$h(k_q)^{-1} = \lambda^{-1/(2q)}\pi^{-1} \begin{cases} k_q\, c_1(k_q), & k_q < 1 \\ \pi c_2(k_q), & k_q \geq 1, \end{cases} \tag{4.40}$$

where $c_1(k_q)$ and $c_2(k_q)$ are known functions that vary slow with k_q, namely

$$\begin{aligned} c_1(k_q) &\coloneqq {}_2\mathcal{F}_1[\{1, 1/(2q)\}; \{1 + 1/(2q)\}, -k_q^{2q}] \in (\pi/4, 1] \\ c_2(k_q) &\coloneqq \widetilde{c}_2 - \frac{{}_2\mathcal{F}_1[\{1, 1 - 1/(2q)\}; \{2 - 1/(2q)\}, -k_q^{-2q}]}{{k_q}^{2q-1}\pi(2q-1)} \in (0.25, 0.5], \end{aligned}$$

with $\widetilde{c}_2 = \mathrm{sinc}\{\pi/(2q)\}^{-1}\pi^{-1}$ and ${}_2\mathcal{F}_1$ denoting hypergeometric series (Abramowitz and Stegun, 1972). For simplicity of notation, in the following sections we write c_1 instead of $c_1(k_q)$ and c_2 instead of $c_2(k_q)$. To obtain representation (4.40) for $k_q < 1$, we used a change of variable, Taylor series expansion and the definition of the hypergeometric series:

$$\int_0^{k_q} \frac{dx}{1 + x^{2q}} = k_q \int_0^1 \frac{dx}{1 + (k_q x)^{2q}} = k_q \int_0^1 \sum_{l=0}^{\infty} (-1)^l (k_q x)^{2ql} = k_q \sum_{l=0}^{\infty} \frac{(-k_q^{2q})^l}{2q\,\{l + 1/(2q)\}}.$$

Similarly, for $k_q \geq 1$, we made use of

$$\begin{aligned} \int_0^{k_q} \frac{dx}{1 + x^{2q}} &= \int_0^{\infty} \frac{dx}{1 + x^{2q}} - \int_0^{k_q^{-1}} \frac{x^{2q-2}dx}{1 + x^{2q}} = \mathrm{sinc}\,\{\pi/(2q)\}^{-1} - \sum_{l=0}^{\infty} \int_0^{k_q^{-1}} (-1)^l x^{2ql+2q-2} \\ &= \mathrm{sinc}\,\{\pi/(2q)\}^{-1} - k_q^{-2q+1} \sum_{l=0}^{\infty} \frac{(-k_q^{-2q})^l}{2q\,\{l + 1 - 1/(2q)\}}. \end{aligned}$$

Formula (4.40) is equivalent to

$$h(k_q) = \begin{cases} K^{-1}\, c_1(k_q)^{-1}, & k_q < 1 \\ \lambda^{1/(2q)} c_2(k_q)^{-1}, & k_q \geq 1. \end{cases} \tag{4.41}$$

Note that for regression spline estimators ($k_q = 0$) the bandwidth $h(0) = K^{-1}c_1(0)^{-1} = K^{-1}$ and for smoothing spline estimators the bandwidth $h(\infty) = \lambda^{1/(2q)}\widetilde{c}_2^{-1}$. Therefore, bandwidth $h(k_q)$ satisfies our requirements.

4.4.2 Asymptotic behavior of the equivalent kernel

In Theorem 3, we show how equivalent kernel $\mathcal{K}(x,t)$ depends on k_q. In particular, k_q determines the transition rate from $\mathcal{K}(x,t)$ to equivalent kernel of regression splines $\mathcal{K}_{rs}(x,t)$ for $k_q < 1$ and to equivalent kernel of smoothing splines $\mathcal{K}_{ss}(x,t)$ for $k_q \geq 1$. More discussion follows after the theorem.

Theorem 3 *The equivalent kernel for spline estimators on $\mathbb{R}$ for $p = 2q-1$ is given by*

$$\begin{cases} c_1\mathcal{K}\left(c_1x, c_1t\right) = \mathcal{K}_{rs}(x,t) - k_q^{2q}\mathcal{K}_1(x,t), & k_q < 1 \\ c_2\mathcal{K}\left(c_2x, c_2t\right) = \mathcal{K}_{ss}(t-x) + k_q^{1-2q}\mathcal{K}_2(x,t), & k_q \geq 1, \end{cases}$$

where $\mathcal{K}_1(x,t)$ and $\mathcal{K}_2(x,t)$ are bounded functions given in the proof, $\mathcal{K}_{rs}(x,t)$ is the regression spline equivalent kernel

$$\mathcal{K}_{rs}(x,t) = \sum_{j=1}^{p}\sum_{l=0}^{2p} \frac{\alpha_l\left(\{x\},\{t\}\right)}{\Pi'_{2p}(r_j)}\, r_j^{|d_{x,t}+l-1|+I_{\{d_{x,t}\leq -l\}}(2p-2)},$$

with α_l, r_j, $d_{x,t}$ defined in Theorem 2 and $\mathcal{K}_{ss}(t-x)$ is the smoothing spline equivalent kernel given in (3.23).

Proof of Theorem 3
From (4.29), we have

$$\mathcal{W}(x,t) = \int_0^K \frac{\phi(u,x)\overline{\phi(u,t)}}{1+\lambda\mu(u)}du = \Re\int_0^{1/2} \frac{2K\phi(Ku,x)\overline{\phi(Ku,t)}}{1+\lambda\mu(Ku)}du.$$

Let first consider $0 \leq k_q < 1$. Scaling $\mathcal{W}(x,t)$ with $c_1^{-1}K^{-1}$, leads to

$$\begin{aligned} c_1\mathcal{K}(c_1x,c_1t) &= \Re\left[\int_0^{1/2} 2\phi(Ku,x/K)\overline{\phi(Ku,t/K)}du\right. \\ &- \left.\int_0^{1/2} \frac{2\lambda\mu(Ku)\phi(Ku,x/K)\overline{\phi(Ku,t/K)}}{1+\lambda\mu(Ku)}du\right] \\ &= \mathcal{K}_{rs}(x,t) - k_q^{2q}\mathcal{K}_1(x,t), \end{aligned}$$

where $\mathcal{K}_{rs}(x,t)$ is the equivalent regression spline kernel on $\mathbb{R}$ and

$$\begin{aligned}
\mathcal{K}_1(x,t) &:= \Re \int_0^{1/2} \frac{2\sin(\pi u)^{2q} Q_{2q-2}(u)\phi(Ku, x/K)\overline{\phi(Ku, t/K)}}{\pi^{2q} Q_{p,M}(u)\{1+\lambda\mu(Ku)\}} du \\
&\leq \frac{2^{2q} Q_{2q-2}(1/2)}{\pi^{2q} Q_{4q-2}(1/2)} \mathcal{K}_{rs}(x,t).
\end{aligned}$$

Using $Q_{lq-2}(1/2) = 2\pi^{lq}(2^{lq}-1)\zeta(lq)$ for the Riemann zeta function $\zeta(lq) = \sum_{i=1}^{\infty} i^{-lq}$, we can get explicit bounds for each q.
For $k_q \geq 1$, we first introduce the following notation.

1. $1 + \lambda\mu(Ku) = \{1 + \lambda(2\pi K u)^{2q}\}\{1 + r_1(u)\}$.
2. $\phi(Ku, x)\overline{\phi(Ku, t)} = \exp\{2\pi \mathtt{i} K u(x-t)\}\{1 + r_2(x,t,u)\}$.
3. $r_q(x,t,u) = \{r_2(x,t,u) - r_1(u)\}\{1 + r_1(u)\}^{-1}$.

Scaling $\mathcal{W}(x,t)$ with $c_2^{-1}\lambda^{1/(2q)}$ and using approximations defined above results in

$$\begin{aligned}
c_2\mathcal{K}(c_2 x, c_2 t) &= \int_{-\infty}^{\infty} \frac{\exp\{2\pi \mathtt{i} u(t-x)\}}{1+(2\pi u)^{2q}} du \\
&+ \Re \int_0^{k_q/2} \frac{2\exp\{2\mathtt{i}u(t-x)\}}{\pi\{1+(2u)^{2q}\}} r_q(x,t,u/k_q) du \\
&- \Re \int_{k_q/2}^{\infty} \frac{2\exp\{2\mathtt{i}u(t-x)\}}{\pi\{1+(2u)^{2q}\}} du = \mathcal{K}_{ss}(x,t) + k_q^{-2q+1}\mathcal{K}_2(x,t),
\end{aligned}$$

where $\mathcal{K}_{ss}(x,t)$ is the smoothing spline kernel on $\mathbb{R}$ and

$$\begin{aligned}
\pi\mathcal{K}_2(x,t) &:= k_q^{2q-1} \Re \int_0^{k_q/2} \frac{2\exp\{2\mathtt{i}u(t-x)\}}{1+(2u)^{2q}} r_q(x,t,u/k_q) du \\
&- \int_1^{\infty} \frac{\cos\{k_q u(t-x)\}}{k_q^{-2q}+u^{2q}} du.
\end{aligned}$$

The second component of $\pi\mathcal{K}_2(x,t)$ is clearly bounded by 1. Now, let us consider $r_q(x,t,u/k_q)$. Since $Q_{4q-2}(u) \leq Q_{p,M}(u) \leq Q_{2q-2}^2(u)$ for odd p, the term $r_1(u)$ is bounded by

$$\frac{(2uk_q)^{2q}}{1+(2uk_q)^{2q}} \left\{ \frac{\operatorname{sinc}(\pi u)^{2q}}{Q_{2q-2}(u)} - 1 \right\} \leq r_1(u) \leq \frac{(2uk_q)^{2q}}{1+(2uk_q)^{2q}} \left\{ \frac{\operatorname{sinc}(\pi u)^{2q} Q_{2q-2}(u)}{Q_{4q-2}(u)} - 1 \right\},$$

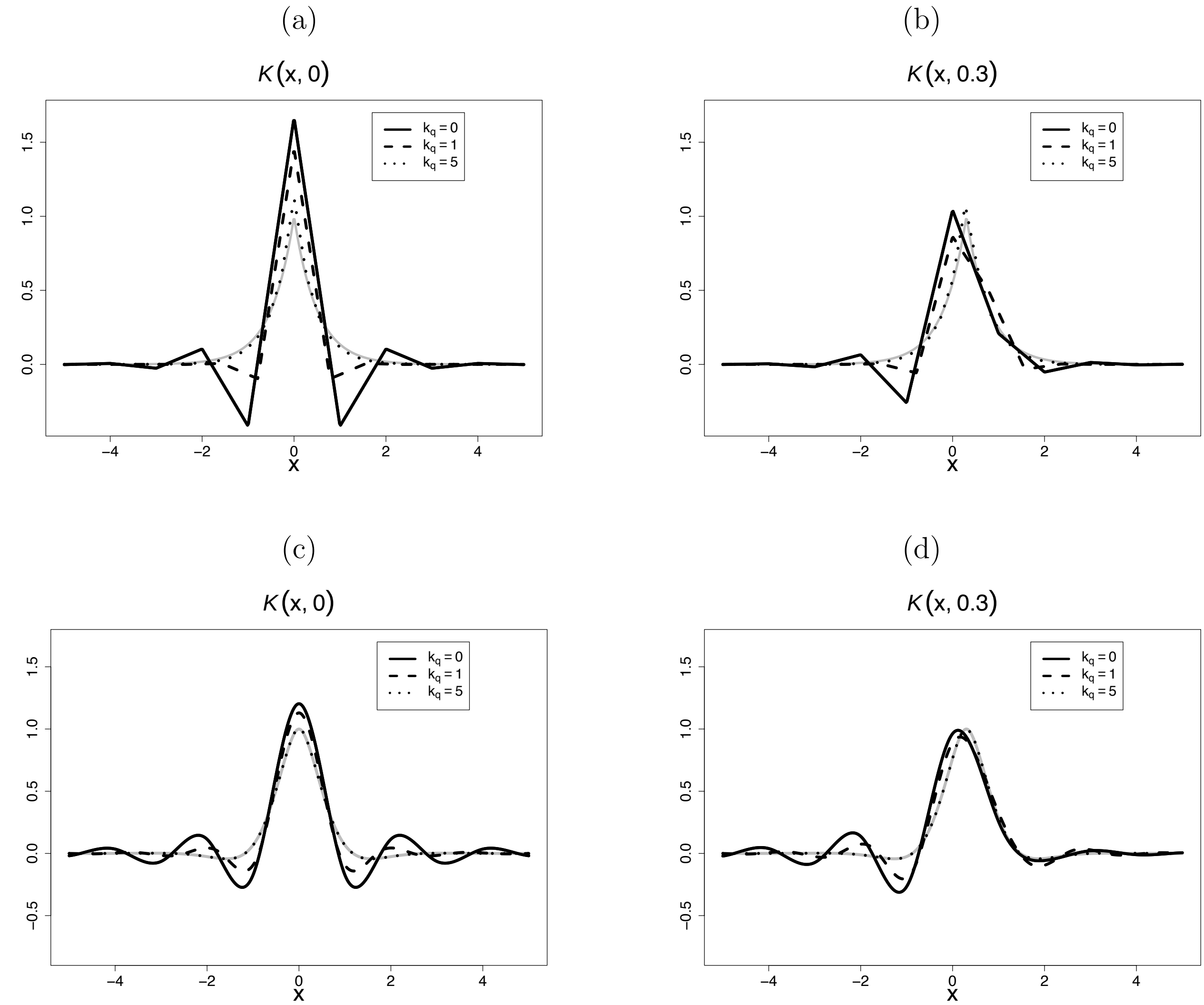

Figure 4.3: Equivalent kernel $\mathcal{K}(x,t)$ for penalized splines for $k_q = 0, 1, 5$ and $M = 5$. (a) $t = 0$, $p = q = 1$, (b) $t = 0.3$, $p = q = 1$, (c) $t = 0$, $p = 2q - 1 = 3$, (d) $t = 0.3$, $p = 2q - 1 = 3$. The grey lines correspond to the smoothing spline kernels.

so that $|r_1(u)| \leq 1 - \operatorname{sinc}(\pi u)^{2q}/Q_{2q-2}(u)$. Observing

$$Q_{2q-2}(u/k_q)\operatorname{sinc}(\pi u/k_q)^{-2q} = 1 + 2\zeta(2q)(u/k_q)^{2q} + O(k_q^{-4q}),$$

we obtain $|r_1(u/k_q)| \leq \zeta(2q)(2k_q)^{-2q} + O(k_q^{-4q})$. Using the same techniques, we obtain $|r_2(x,t,u/k_q)| \leq 8\zeta(2q)(2k_q)^{-2q} + O(k_q^{-4q})$. In principle, one can also find a lower bound

for $1+r_1(u/k_q)$ depending on q and k_q, but it is enough to note that $1+r_1(u/k_q) \geq 1/2$. Finally, $|r_q(x,t,u/k_q)| \leq 18\zeta(2q)(2k_q)^{-2q} + O(k_q^{-4q})$ and hence, the first term in $\pi\mathcal{K}_2(x,t)$ is also bounded for any $k_q \geq 1$.
It remains to set $\mathcal{K}_{rs}(x,t)$. $\mathcal{K}_{rs}(x,t)$ is obtained from (4.36), scaling $\mathcal{W}(x,t)$ with K and setting $P_{2p}(u) = \Pi_{2p}(u)$. $\square$

From Theorem 3, it follows that $\lim_{k_q\to\infty} c_2\mathcal{K}(c_2x, c_2t) = \widetilde{c}_2^{-1}\mathcal{K}_{ss}\{(x-t)/\widetilde{c}_2\}$ and $\lim_{k_q\to 0} c_1\mathcal{K}(c_1x, c_1t) = \mathcal{K}_{rs}(x,t)$, that is, $\mathcal{K}(x,t)$ varies smoothly between $\mathcal{K}_{rs}(x,t)$ and $\mathcal{K}_{ss}(x,t)$, scaled with appropriate constants. This is visualized in Figures 4.3. Figure 4.3 depicts the penalized spline kernel $\mathcal{K}(x,t)$ for $M=5$ at $t=0$ and $t=0.3$ for different values of k_q and for $p=1,3$. As k_q grows, $\mathcal{K}(x,t)$ becomes more symmetric and for $k_q=5$ is already non-distinguishable from the smoothing spline kernel shown in grey.

4.4.3 Moments and exponential decay of the equivalent kernel

The next two lemmas are technical and are used in the proof of the Theorem 4 on the pointwise behavior of periodic splines in Section 4.4.4. As Nychka (1995) notices, one does not need to know the exact form of the kernel in order to study the pointwise bias and variance of smoothing splines. Similarly, in the proof of our Theorem 4, it turns out that only the following two lemmas are crucial.

Lemma 15 *For $x,t\in\mathbb{R}$, it holds for $0\leq m\leq \min(p, 2q-1)$ that*

$$\int_{-\infty}^{\infty}(t-x)^m\,\mathcal{W}(x,t)\,dt = \int_0^1 (t-x)^m\,W(x,t)\,dt = \delta_{m,0},$$

while for $m=\min(p+1,2q)$,

$$\begin{aligned}\int_{-\infty}^{\infty}(t-x)^m\,\mathcal{W}(x,t)\,dt &= \int_0^1 (t-x)^m\,W(x,t)\,dt \\ &= I_{\{p\leq 2q-1\}}\left\{\frac{2\mathcal{B}_{p+1}\left(\left\{\frac{p+1}{2}\right\}\right)}{N^{p+1}} - \frac{\mathcal{B}_{p+1}\left(\left\{Kx+\frac{p+1}{2}\right\}\right)}{K^{p+1}}\right\} - I_{\{p\geq 2q-1\}}\frac{\lambda\,(2q)!}{(-1)^q},\end{aligned}$$

Proof of Lemma 15
From (4.33), (4.34) and the fact that $\mathcal{W}(x,t)$ is a real-valued function it follows that

$$\mathcal{W}(x,t) = \int_{-\infty}^{\infty} \overline{a(u,x)} \exp(-2\pi i t u)\, du,$$

with $a(u,x)$ defined as

$$\begin{aligned} a(u,x) \;&\coloneqq\; \frac{\operatorname{sinc}\{\pi(u/K)\}^{p+1}\phi(u,x)}{\sqrt{Q_{p,M}(u/K)}\{1+\lambda\mu(u)\}} = \int_{-\infty}^{\infty} \mathcal{W}(x,t)\exp(-2\pi \mathtt{i} t u)dt \\ &= \int_0^1 W(x,t)\exp(-2\pi \mathtt{i} t u)\, dt, \;\; u \in \mathbb{R}, \end{aligned} \tag{4.42}$$

where the last equality is obtained using (4.30) and the Poisson summation formula. Note that for $u \in \mathbb{Z}$, function $a(u,x)$ coincides with the u-th Fourier coefficient of $W(x,t)$ given in (4.31). Properties of the Fourier transform and (4.42) ensure that

$$\begin{aligned} &\int_{-\infty}^{\infty} (2\pi \mathtt{i} t)^m \mathcal{W}(x,t)\exp(2\pi \mathtt{i} t u)dt = \frac{\partial^m}{\partial u^m}\int_{-\infty}^{\infty} \mathcal{W}(x,t)\exp(2\pi \mathtt{i} t u)dt \\ =&\frac{\partial^m}{\partial u^m}\int_0^1 W(x,t)\exp(2\pi \mathtt{i} t u)dt = \int_0^1 (2\pi \mathtt{i} t)^{2q} W(x,t)\exp(2\pi \mathtt{i} t u)dt = \frac{\partial^m}{\partial u^m}\overline{a(u,x)}. \end{aligned}$$

Evaluating derivative of $a(u,x)$ at $u=0$ and grouping the terms of $a(u,x)$ in (4.42), we represent

$$\int_{-\infty}^{\infty} (2\pi \mathtt{i} t)^m \mathcal{W}(x,t)dt = I_1 + I_2 + I_3, \tag{4.43}$$

where

$$\begin{aligned} I_1 \coloneqq& \frac{\partial^m}{\partial u^m}\left[\exp(2\pi \mathtt{i} x u)\left\{1 + \frac{\operatorname{sinc}(\pi u/K)^{2p+2} - Q_{p,M}(u/K)}{Q_{p,M}(u/K)}\right\}\right]_{u=0} \\ I_2 \coloneqq& \frac{\partial^m}{\partial u^m}\left[\frac{\{\sin(\pi u/K)\operatorname{sinc}(\pi u/K)\}^{p+1}}{Q_{p,M}(u/K)}\sum_{l\neq 0}\frac{\exp\{2\pi \mathtt{i} x(u+lK)\}}{\left\{(-1)^l \pi(u/K+l)\right\}^{p+1}}\right]_{u=0} \\ I_3 \coloneqq& \frac{\partial^m}{\partial u^m}\left[-\frac{\lambda\mu(u)\operatorname{sinc}(\pi u/K)^{p+1}\phi(u,x)}{\sqrt{Q_{p,M}(u/K)}\{1+\lambda\mu(u)\}}\right]_{u=0}. \end{aligned}$$

The idea is to represent each of these components as a product of $\sin(\pi u/K)^n, n \in \mathbb{Z}$ and some function that is differentiable at 0. After that, we use that $Q_{p,M}(0) = \mu(0) = \phi(0,x) = 1$,

$$\frac{\partial^m}{\partial u^m}\{\sin(\pi u/K)^n\}_{u=0} = \begin{cases} 0, & m = 0, \dots n-1 \\ n!\,(\pi/K)^n, & m = n \end{cases}$$

and the Fourier series of the periodic Bernoulli polynomials

$$\mathcal{B}_{p+1}(\{x\}) = (-1)^p (p+1)! \sum_{s \neq 0} \exp(-2\pi \mathtt{i} sx)/(2\pi \mathtt{i} s)^{p+1}.$$

To handle I_1, we regroup series representation (4.12) of $Q_{p,M}$ as

$$\begin{aligned} Q_{p,M}(u/K) = \operatorname{sinc}(\pi u/K)^{2p+2} &+ 2\{\sin(\pi u/K)\operatorname{sinc}(\pi u/K)\}^{p+1} \sum_{l\neq 0} \frac{(-1)^{lM(p+1)}}{\{\pi(u/K + lM)\}^{p+1}} \\ &+ \frac{\sin(\pi u/K)^{2p+2}}{\pi^{2p+2}} \sum_{j=1}^{M-1} \left\{ \sum_{l=-\infty}^{\infty} \frac{(-1)^{(j+lM)(p+1)}}{(u/K + j + lM)^{p+1}} \right\}^2 \\ &+ \frac{\sin(\pi u/K)^{2p+2}}{\pi^{2p+2}} \left\{ \sum_{l \neq 0} \frac{(-1)^{lM(p+1)}}{(u/K + lM)^{p+1}} \right\}^2. \end{aligned}$$

With this, the wanted representation for the components of I_1 is obtained. Putting it all together, we find

$$I_1 = \begin{cases} (2\pi \mathtt{i} x)^m, & m = 0, \dots p \\ (2\pi \mathtt{i} x)^{p+1} + 2(2\pi \mathtt{i}/N)^{p+1} \mathcal{B}_{p+1}\left(\left\{\frac{p+1}{2}\right\}\right), & m = p+1. \end{cases}$$

The components of I_2 are already represented as a product of $\sin(\pi u/K)^{p+1}$ and a function that is differentiable at 0. Thus, we find that

$$I_2 = \begin{cases} 0, & m = 0, \dots p \\ -(2\pi \mathtt{i}/K)^{p+1} \mathcal{B}_{p+1}\left(\left\{Kx + \frac{p+1}{2}\right\}\right), & m = p+1. \end{cases}$$

To find I_3, we use $\mu(u) = (2K)^{2q} \sin(\pi u/K)^{2q} Q_{2p-2q}(u/K)/Q_{p,M}(u/K)$ that follows from (4.19). We obtain

$$I_3 = \begin{cases} 0, & m = 0, \ldots 2q-1 \\ -\lambda\,(2\pi)^{2q}\,(2q)!, & m = 2q. \end{cases}$$

To obtain integrals $\int_{-\infty}^{\infty} (t-x)^m\, \mathcal{W}(x,t)\, dt = \int_0^1 (t-x)^m\, W(x,t)\, dt$, one needs to expand $(t-x)^m$ and use (4.43). □

The next lemma deals with the decay of equivalent kernel $\mathcal{K}(x,t)$.

Lemma 16 *Kernel $\mathcal{K}(x,t)$, $x,t \in \mathbb{R}$ decays exponentially, i.e. there are constants $0 < C < \infty$ and $0 < \gamma < 1$ such that*

$$|\mathcal{K}(x,t)| < C\,\gamma^{|x-t|}. \tag{4.44}$$

Proof of Lemma 16

Since $\mathcal{K}(x,t)$ is defined as function $\mathcal{W}(x,t)$ scaled with $h(k_q)$, from (4.36) and (4.41), one finds for $k_q < 1$ that

$$c_1\mathcal{K}(c_1 x, c_1 t) = \sum_{j=1}^{p}\sum_{l=0}^{2p} \frac{\alpha_l\left(\{x\},\{t\}\right)}{P'_{2p}(r_j)} r_j^{|\lfloor x\rfloor - \lfloor t\rfloor + l - 1| + I_{\{\lfloor x\rfloor - \lfloor t\rfloor \le -l\}}(2p-2)}, \tag{4.45}$$

while for $k_q \geq 1$,

$$\begin{aligned} \pi c_2\mathcal{K}(\pi c_2 x, \pi c_2 t) &= k_q \sum_{j=1}^{p}\sum_{l=0}^{2p} \frac{\alpha_l\left(\{x k_q\},\{t k_q\}\right)}{P'_{2p}(r_j)} \\ &\times\; r_j^{|\lfloor x k_q\rfloor - \lfloor t k_q\rfloor + l - 1| + I_{\{\lfloor x k_q\rfloor - \lfloor t k_q\rfloor \le -l\}}(2p-2)}. \end{aligned} \tag{4.46}$$

Here, polynomial P_{2p} is given in (4.35) and the $r_j = r_j(k_q)$ are its roots with $|r_j| < 1$. If k_q is a bounded constant, then $r_j = r_j(k_q) \nrightarrow \exp(-2\pi \mathtt{i} u),\ u \in (0,1)$ since

$$P_{2p}\left\{\exp(-2\pi\mathtt{i}u)\right\} = \exp(-2\pi\mathtt{i}pu)\left\{Q_{p,M}(u) + (2k_q/\pi)^{2q}\sin(\pi u)^{2q} Q_{2p-2q}(u)\right\} \neq 0,$$

where the relationship follows from formula (4.37). Similarly, $r_j = r_j(k_q) \nrightarrow 0$ and $0 < \gamma < 1$ can be defined as follows.

$$\gamma := \begin{cases} \sup_{j,k_q} |r_j(k_q)|, & k_q < 1 \\ \sup_{j,k_q} \left|r_j(k_q)^{k_q}\right|, & 1 \le k_q < \infty, \end{cases}$$

while

$$C := \sup_{k_q,j} \frac{p(2p+1)\sup_{l,x,t}\alpha_l(\{x\},\{t\})}{\left|P'_{2p}\{r_j(k_q)\}\right| |r_j(k_q)|^{l+1}} < \infty.$$

For $k_q \to \infty$, we know from Theorem 3 that

$$\lim_{k_q\to\infty} \mathcal{K}(x,t) = \mathcal{K}_{ss}\left\{(x-t)/\widetilde{c}_2\right\}/\widetilde{c}_2.$$

To obtain the bound on the smoothing spline kernel $\mathcal{K}_{ss}(x)$, the expression given in Theorem 3 can be rewritten as

$$\begin{aligned} |\mathcal{K}_{ss}(x-t)| &= \Bigg| -I_{\{q \text{ is odd}\}}\frac{\exp(-|x-t|)}{2q} + \sum_{j=0}^{\lfloor (q-1)/2 \rfloor} \frac{\exp\left[-|x-t|\sin\{\pi(2j+1)/(2q)\}\right]}{q} \\ &\times \sin\left[\frac{\pi(2q-1)(2j+1)}{2q} - |x-t|\cos\left\{\frac{\pi(2j+1)}{2q}\right\}\right]\Bigg| \\ &\le \frac{q+1}{2q}\exp\left\{-|x-t|\sin(\pi/2q)\right\}, \end{aligned}$$

so that one can set $\gamma := \exp\left[-\sin\{\pi/(2q)\}/\widetilde{c}_2\right] \in (0,1)$ and $C := (q+1)/(2q\widetilde{c}_2) < \infty$ for $k_q \to \infty$. □

4.4.4 Pointwise bias and variance of spline estimators

In this section, we obtain new results concerning the pointwise bias and pointwise variance of the periodic spline estimator $\widehat{f}$. The local asymptotic behavior is given in terms of the universal bandwidth $h(k_q)$.

Theorem 4 *Let the model (2.1) hold and $\widehat{f}(x) \in S_{\mathtt{per}}(2q-1; \underline{\tau}_K)$ be the solution to (4.1) with $\{x_i = i/N\}_{i=1}^N$ and $\underline{\tau}_K = \{i/K\}_{i=0}^K$. Then, for $f \in C_{\mathtt{per}}^{p+1}$ such that $f^{(2q)} \in$*

$C^{0,\alpha}([0,1])$, i.e. $f^{(2q)}$ is Hölder continuous with $|f^{(2q)}(x) - f^{(2q)}(t)| \leq L|x-t|^\alpha$, $\forall x, t \in [0,1]$ and $\alpha \in (0,1]$, it holds at any $x \in [0,1]$

$$\begin{aligned}
\mathrm{E}\left\{\widehat{f}(x)\right\} - f(x) &= -h(k_q)^{2q}\frac{f^{(2q)}(x)}{(2q)!}C(k_q,x) + o\left\{h(k_q)^{2q}\right\} \\
\mathrm{Var}\left\{\widehat{f}(x)\right\} &= \frac{\sigma^2}{Nh(k_q)}\int_{-\infty}^{\infty}\mathcal{K}^2\left\{x/h(k_q),t\right\}dt + o\left\{N^{-1}h(k_q)^{-1}\right\},
\end{aligned}$$

where

$$C(k_q,x) \coloneqq \begin{cases} c_1^{2q}\left[\mathcal{B}_{2q}\left(\{Kx\}\right) + (-1)^q(2q)!\pi^{-2q}k_q^{2q} - \mathcal{B}_{2q}M^{-2q}\right], & k_q < 1 \\ c_2^{2q}\left[(-1)^q(2q)! + \mathcal{B}_{2q}\left(\{Kx\}\right)\pi^{2q}k_q^{-2q} - 2\mathcal{B}_{2q}\pi^{2q}(k_qM)^{-2q}\right], & k_q \geq 1 \end{cases}$$

and $\int_{-\infty}^{\infty}\mathcal{K}^2\left\{x/h(k_q),t\right\}dt < C^2/\log(\gamma^{-1})$ for some $C \in (0,\infty)$ and $\gamma \in (0,1)$, both depending on k_q, explicitly given in the proof of Lemma 16.

Proof of Theorem 4

The proof is based on the kernel regression methods, discussed in Section 3.1.1. To show that the remainder terms are asymptotically negligible, we use the exponential decay (4.44) and methods similar to Huang and Studden (1993). From (4.27),

$$\begin{aligned}
\mathrm{E}\left\{\widehat{f}(x)\right\} &= \mathrm{E}\left\{\frac{1}{N}\sum_{i=1}^{N}W(x,i/N)Y_i\right\} = \int_0^1 W(x,t)f(t)dt + O\left(N^{-1}\right) \\
&= \int_{-\infty}^{\infty}\mathcal{W}\left(x,t\right)f(t)dt + O\left(N^{-1}\right).
\end{aligned}$$

The Taylor series of f around x at point t are

$$f(t) = f(x) + f'(x)(x-t) + \ldots + \frac{f^{(2q-1)}(x)}{(2q-1)!}(x-t)^{2q-1} + \frac{f^{(2q)}(\xi_{x,t})}{(2q)!}(x-t)^{2q},$$

where $\xi_{x,t}$ is a point between x and t. Using this and Lemma 15 results in

$$\begin{aligned}
\mathrm{E}\left\{\widehat{f}(x)\right\} - f(x) &= \int_{-\infty}^{\infty} \mathcal{W}(x,t)\,(x-t)^{2q} \frac{f^{(2q)}(\xi_{x,t})}{(2q)!} dt + O(N^{-1}) \\
&= \frac{f^{(2q)}(x)}{(2q)!} \int_{-\infty}^{\infty} \mathcal{W}(x,t)\,(x-t)^{2q} dt + R_\xi(x) + O(N^{-1}) \\
&= h(k_q)^{2q} \frac{f^{(2q)}(x)}{(2q)!} \int_{-\infty}^{\infty} \mathcal{K}(x_h,t_h)\,(x_h-t_h)^{2q}\, dt_h + R_\xi(x) + O(N^{-1}),
\end{aligned}$$

where $x_h \coloneqq x/h(k_q)$, $t_h \coloneqq t/h(k_q)$ and

$$R_\xi(x) \coloneqq h(k_q)^{2q} \int_{-\infty}^{\infty} \mathcal{K}(x_h,t_h)\,(x_h-t_h)^{2q} \frac{f^{(2q)}(\xi_{x,t}) - f^{(2q)}(x)}{h(k_q)(2q)!} dt.$$

From Lemma 15, $\int_{-\infty}^{\infty} \mathcal{K}(x_h,t)\,(x_h-t)^{2q}\, dt = -C(k_q,x)$, where $C(k_q,x)$ is given in Theorem 4. It remains to show that $R_\xi(x) = o\left\{h(k_q)^{2q}\right\}$. Using techniques similar to Huang and Studden (1993),

$$\begin{aligned}
R_\xi(x) &= h(k_q)^{2q} \sum_{l=-\infty}^{\infty} \int_{x+(l-1)h}^{x+lh} \mathcal{K}(x_h,t_h)\,(x_h-t_h)^{2q} \frac{f^{(2q)}(\xi_{x,t}) - f^{(2q)}(x)}{h(k_q)(2q)!} dt \\
&\leq h(k_q)^{2q+\alpha} CL \sum_{l=-\infty}^{\infty} \int_{x+(l-1)h}^{x+lh} \gamma^{|x_h-t_h|} \frac{|x_h-t_h|^{2q+\alpha}}{h(k_q)(2q)!} dt \\
&\leq h(k_q)^{2q+\alpha} \frac{2CL}{(2q)!} \sum_{l=1}^{\infty} \gamma^{l-1} l^{2q+\alpha} = o\left\{h(k_q)^{2q}\right\},
\end{aligned}$$

where the exponential bound on the kernel from Lemma 16 together with the Hölder continuity of $f^{(2q)}$ have been used.
Next, the variance of $\widehat{f}(x)$ is given by

$$\mathrm{Var}\left\{\widehat{f}(x)\right\} = \frac{\sigma^2}{N^2} \sum_{i=1}^{N} W^2(x,i/N) = \frac{\sigma^2}{N} \int_0^1 W^2(x,t)dt + O(N^{-1}).$$

Let us define $K(x,t)$ via

$$h(k_q)^{-1} K(x_h,t_h) = W(x,t) = \sum_{l=-\infty}^{\infty} \mathcal{W}(x,t+l) = h(k_q)^{-1} \sum_{l=-\infty}^{\infty} \mathcal{K}(x_h,t_h+l_h),$$

for $l_h := l/h(k_q)$. Then, using periodicity of $W(x,t)$ and a change of variable t_h in the integration on the last step, we have:

$$\begin{aligned}\int_0^1 W^2(x,t)dt &= \int_{x-\frac{1}{2}}^{x+\frac{1}{2}} W^2(x,t)dt = \frac{1}{h(k_q)^2}\int_{x-\frac{1}{2}}^{x+\frac{1}{2}} K^2(x_h,t_h)dt \\ &= \frac{1}{h(k_q)}\left\{\int_{-\infty}^{\infty}\mathcal{K}^2(x_h,t)dt + R_k(x)\right\},\end{aligned}$$

for

$$\begin{aligned}h(k_q)R_k(x) &= \int_{x-\frac{1}{2}}^{x+\frac{1}{2}} K^2(x_h,t_h)dt - \int_{-\infty}^{\infty}\mathcal{K}^2(x_h,t_h)dt \\ &= \int_{x-\frac{1}{2}}^{x+\frac{1}{2}}\left\{K^2(x_h,t_h)-\mathcal{K}^2(x_h,t_h)\right\}dt - \int_{-\infty}^{x-\frac{1}{2}}\mathcal{K}^2(x_h,t_h)dt - \int_{x+\frac{1}{2}}^{\infty}\mathcal{K}^2(x_h,t_h)dt.\end{aligned}$$

Now, we can make use of $K(x_h,t_h) = \sum_{l=-\infty}^{\infty}\mathcal{K}(x_h,t_h+l_h)$ and of the exponential decay of $\mathcal{K}(x,t)$ found in Lemma 16 to bound terms in $h(k_q)R_k(x)$. That is,

$$\begin{aligned}&\int_{x-\frac{1}{2}}^{x+\frac{1}{2}}\left\{K^2(x_h,t_h)-\mathcal{K}^2(x_h,t_h)\right\}dt \\ =\ &\int_{x-\frac{1}{2}}^{x+\frac{1}{2}}\sum_{l\neq 0}\mathcal{K}(x_h,t_h+l_h)\left\{\sum_{l\neq 0}\mathcal{K}(x_h,t_h+l_h)+2\mathcal{K}(x_h,t_h)\right\}dt \\ \leq\ & C^2\int_{x-\frac{1}{2}}^{x+\frac{1}{2}}\sum_{l\neq 0}\gamma^{|x_h-t_h-l_h|}\left(\sum_{l\neq 0}\gamma^{|x_h-t_h-l_h|}+2\gamma^{|x_h-t_h|}\right)dt \\ =\ & 2C^2\int_0^{\frac{1}{2}}\sum_{l\neq 0}\gamma^{|t_h+l_h|}\left(\sum_{l\neq 0}\gamma^{|t_h+l_h|}+2\gamma^{t_h}\right)dt\end{aligned}$$

To compute the bound we use the sum under the integral

$$\sum_{l\neq 0}\gamma^{|t_h+l_h|} = \left(\gamma^{t_h}+\gamma^{-t_h}\right)\gamma^{1/h(k_q)}/\left\{1-\gamma^{1/h(k_q)}\right\},\ t\in[0,1/2].$$

Hence

$$\begin{aligned}
&\int_{x-\frac{1}{2}}^{x+\frac{1}{2}} \left\{K^2(x_h,t_h) - \mathcal{K}^2(x_h,t_h)\right\} dt \\
\leq & 4C^2 \int_0^{\frac{1}{2}} \left[\frac{\left(\gamma^{t_h}+\gamma^{-t_h}\right)^2 \gamma^{2/h(k_q)}}{2\left\{1-\gamma^{1/h(k_q)}\right\}^2} + \frac{\left(\gamma^{t_h}+\gamma^{-t_h}\right)\gamma^{|t_h|+h(k_q)^{-1}}}{1-\gamma^{1/h(k_q)}} \right] dt \\
= & h(k_q) \frac{C^2\gamma^{1/h(k_q)}\left\{3-4\gamma^{1/h(k_q)}+\gamma^{2/h(k_q)}+2h(k_q)^{-1}\log\left(\gamma^{-1}\right)\right\}}{\left\{1-\gamma^{1/h(k_q)}\right\}\log\left(\gamma^{-1}\right)} \\
\leq & h(k_q) \frac{2C^2\gamma^{1/h(k_q)}\left\{2+h(k_q)^{-1}\log\left(\gamma^{-1}\right)\right\}}{\left\{1-\gamma^{1/h(k_q)}\right\}\log\left(\gamma^{-1}\right)}.
\end{aligned}$$

Also,

$$\begin{aligned}
\int_{-\infty}^{x-\frac{1}{2}} \mathcal{K}^2(x_h,t_h)dt + \int_{x+\frac{1}{2}}^{\infty} \mathcal{K}^2(x_h,t_h)dt &\leq C^2 \left\{ \int_{-\infty}^{x-\frac{1}{2}} \gamma^{2(x_h-t_h)}dt + \int_{x+\frac{1}{2}}^{\infty} \gamma^{2(t_h-x_h)}dt \right\} \\
&= h(k_q)\frac{C^2\gamma^{1/h(k_q)}}{\log(\gamma^{-1})}.
\end{aligned}$$

In a similar fashion, one finds $\int_{-\infty}^{\infty} \mathcal{K}^2(x_h,t)dt \leq C^2/\log(\gamma^{-1})$. Putting it all together gives

$$|R_k(x)| \leq \frac{C^2\gamma^{1/h(k_q)}}{\log(\gamma^{-1})}\left[1+\frac{4+2h(k_q)\log(\gamma^{-1})}{\left\{\gamma^{1/h(k_q)}-1\right\}^2}\right] = O\left\{h(k_q)^{-1}\gamma^{1/h(k_q)}\right\} = o(1),$$

proving the theorem. □

The results on the local asymptotic properties of periodic spline estimators allow us to make the bounds for the choice of optimal λ and K, given k_q, more precise.

Lemma 17 *Under the assumptions of Theorem 4,*

$$\begin{aligned}
IMSE(\widehat{f}) &= \frac{h(k_q)^{4q}}{(2q)!^2}\|f^{(2q)}C(k_q)\|^2 + o\left\{h(k_q)^{4q}\right\} \\
&+ \frac{\sigma^2 c_3}{Nh(k_q)}\int_0^1\int_{-\infty}^{\infty}\mathcal{K}^2(x,t)\,dtdx + o\left\{N^{-1}h(k_q)^{-1}\right\},
\end{aligned}$$

where $\left\|f^{(2q)}C(k_q)\right\| = \left[\int_0^1 \left\{f^{(2q)}(x)C(k_q,x)\right\}^2 dx\right]^{1/2}$, $c_3(k_q) = I_{\{k_q\geq 1\}} + c_1 I_{\{k_q<1\}}$. *The* *asymptotic optimal bandwidth depending on k_q is*

$$h_{opt}(k_q) \coloneqq C_{opt}(k_q) c_3 N^{-1/(4q+1)},$$

where

$$C_{opt}(k_q) \coloneqq \frac{\sigma^2 (2q)! \int_0^1 \int_{-\infty}^{\infty} \mathcal{K}^2\left(x,t\right) dtdx}{4q \left\|f^{(2q)} C(k_q)\right\|^2}.$$

For $k_q < 1$, the asymptotic optimal K is $K_{opt} \coloneqq c_1^{-2} C_{opt}^{-1}(k_q) N^{1/(4q+1)}$ (achieved for $\lambda = k_q^{2q} \left\{c_1^2 C_{opt}(k_q)/\pi\right\}^{2q} N^{-2q/(4q+1)}$). For $k_q \geq 1$, the asymptotic optimal λ is $\lambda_{opt} \coloneqq c_2^{2q} C_{opt}(k_q)^{2q} N^{-2q(4q+1)}$ (achieved for $K = k_q \left\{\pi c_2 C_{opt}(k_q)\right\}^{-1} N^{1/(4q+1)}$).

Proof of Lemma 17

We tackle the proof in three steps.

First, integrating the squared pointwise bias and variance given in Theorem 4, we obtain

$$\begin{aligned} IMSE(\widehat{f}) &= \frac{h(k_q)^{4q}}{(2q)!^2} \|f^{(2q)} C(k_q)\|^2 + o\left\{h(k_q)^{4q}\right\} \\ &+ \frac{\sigma^2}{N h(k_q)} \int_0^1 \int_{-\infty}^{\infty} \mathcal{K}^2\left\{x/h(k_q), t\right\} dtdx + o\left\{N^{-1} h(k_q)^{-1}\right\}, \end{aligned}$$

Second, we show that $\int_0^1 \int_{-\infty}^{\infty} \mathcal{K}^2\left\{x/h(k_q), t\right\} dtdx$ depends on k_q and q only, and hence we can treat it as a constant while minimizing the IMSE over $h(k_q)$ for fixed k_q. Using (4.45) and expression for bandwidth (4.41), we obtain for $k_q < 1$

$$\begin{aligned} \mathcal{K}\{x/h(k_q), t\} &= \sum_{j=1}^{2q-1} \sum_{l=0}^{4q-2} \frac{\alpha_l\left(\{Kx\}, \{c_1^{-1} t\}\right)}{c_1 P'_{4q-2}(r_j)} r_j^{\left|\lfloor Kx \rfloor - \lfloor c_1^{-1} t \rfloor + l - 1\right| + I_{\{\lfloor Kx \rfloor - \lfloor c_1^{-1} t \rfloor \leq -l\}}(4q-4)} \\ &= g(\{Kx\}, \{c_1^{-1} t\}, \lfloor Kx \rfloor - \lfloor c_1^{-1} t \rfloor, k_q, q), \end{aligned} \tag{4.47}$$

where $\alpha_l(x,t)$, $l = 0, \ldots, 4q-2$ are given in Theorem 2 and for fixed x,t depend on q only; $\{r_j\}_{j=0}^{2q-1}$ and $P'_{4q-2}(r_j)$ depend on k_q and q; function g is defined as follows:

$$g(x, t, y, k_q, q) \coloneqq \sum_{j=1}^{2q-1} \sum_{l=0}^{4q-2} \frac{\alpha_l\left(x, t\right)}{c_1 P'_{4q-2}(r_j)} r_j^{|y+l-1| + I_{\{y \leq -l\}}(4q-4)}.$$

Further, with change of variable $c^{-1}t - \lfloor Kx \rfloor$, we obtain

$$\int_{-\infty}^{\infty} g^2\left(\{Kx\}, \{c_1^{-1}t\}, \lfloor Kx \rfloor - \left\lfloor c_1^{-1}t \right\rfloor, k_q, q\right) dt = c_1 \int_{-\infty}^{\infty} g^2\left(\{Kx\}, \{y\}, -\lfloor t \rfloor, k_q, q\right) dt.$$

Change of variable Kx and simple computations lead to

$$\int_0^1 \int_{-\infty}^{\infty} g^2\left(\{Kx\}, \{y\}, -\lfloor t \rfloor, k_q, q\right) dt dx = \int_0^1 \int_{-\infty}^{\infty} g^2\left(x, \{y\}, -\lfloor t \rfloor, k_q, q\right) dt dx.$$

Hence, for $k_q < 1$ it holds that

$$\int_0^1 \int_{-\infty}^{\infty} \mathcal{K}^2\left\{x/h(k_q), t\right\} dt dx = c_1 \int_0^1 \int_{-\infty}^{\infty} \mathcal{K}^2\left(x, t\right) dt dx,$$

where $\mathcal{K}(x,t)$ depends on k_q and q only for fixed x, t. For $k_q > 1$, it follows from (4.46) and (4.41) that

$$\begin{aligned} \mathcal{K}\{x/h(k_q), t\} &= \widetilde{k}_q \sum_{j=1}^{2q-1} \sum_{l=0}^{4q-2} \frac{\alpha_l\left(\{Kx\}, \{\widetilde{k}_q t\}\right)}{P'_{4q-2}(r_j)} r_j^{\left|\lfloor Kx \rfloor - \left\lfloor \widetilde{k}_q t \right\rfloor + l - 1\right| + I_{\{\lfloor Kx \rfloor - \lfloor \widetilde{k}_q t \rfloor \le -l\}}(4q-4)} \\ &= \widetilde{k}_q g(\{Kx\}, \{\widetilde{k}_q t\}, \lfloor Kx \rfloor - \lfloor \widetilde{k}_q t \rfloor, k_q, q) \end{aligned} \tag{4.48}$$

with $\widetilde{k}_q := c_2^{-1} k_q / \pi$. We continue in a similar fashion as before to obtain for $k_q \ge 1$

$$\int_0^1 \int_{-\infty}^{\infty} \mathcal{K}^2\left\{x/h(k_q), t\right\} dt dx = \int_0^1 \int_{-\infty}^{\infty} \mathcal{K}^2\left(x, t\right) dt dx.$$

Third, we show that $\|f^{(2q)} C(k_q)\|^2$ is asymptotically a constant with respect to $h(k_q)$ if k_q is fixed. For this, we will need

$$\begin{aligned} \int_0^1 \left\{f^{(2q)}(x)\right\}^2 \mathcal{B}_{2q}(\{Kx\})^2 dx &= \int_0^1 K^{-1} \sum_{j=0}^{K-1} \left\{f^{(2q)}\left(\frac{x+j}{K}\right)\right\}^2 \mathcal{B}_{2q}(x)^2 dx \\ &= \left[\int_0^1 \left\{f^{(2q)}(t)\right\}^2 dt + O(K^{-1})\right] \int_0^1 \mathcal{B}_{2q}^2(x) dx. \end{aligned}$$

Finally, we obtain the asymptotic optimal bandwidth by minimizing over $h(k_q)$ the asymptotic IMSE

$$\frac{h(k_q)^{4q}}{(2q)!^2}\|f^{(2q)}C(k_q)\|^2 + \frac{\sigma^2 c_3}{Nh(k_q)}\int_0^1\int_{-\infty}^{\infty}\mathcal{K}^2(x,t)\,dtdx.$$

For finding asymptotic optimal K for $k_q < 1$ we use $h(k_q) = c_1^{-1}K^{-1}$. For finding asymptotic optimal λ for $k_q \geq 1$, we use for $h(k_q) = c_2^{-1}\lambda^{1/(2q)}$. $\square$

Compared to Corollary 1, the results of Lemma 17 are more precise.

5 Discussion

In this dissertation, the local asymptotic properties of periodic spline estimators are obtained in Theorem 4 by employing the equivalent kernel $\mathcal{W}(x,t)$ with bandwidth $h(k_q)$. As can be deduced from the proof of Theorem 4, the obtained pointwise bias and variance of periodic splines coincide in the interior with those of non-periodic splines. The asymptotic properties of periodic splines depend on the bandwidth that behaves differently depending on k_q. Crucial parameter k_q determines the type of the global and local asymptotics, the pointwise bias, the constants in the bandwidth and the shape of the equivalent kernel. Unlike the standard approaches on studying the asymptotic properties of spline-based estimators, our approach does not require an approximation of the spline estimator. Instead, we are able to work with the exact weight function of the (periodic) spline estimator obtaining precise results.

In this chapter, we consider possible generalizations of our results.

The assumption of equidistant knots and observations dominates in the literature on the equivalent kernels. However, we can generalize the results following Huang and Studden (1993). In particular, if the design points $\{x_i\}_{i=1}^N$ have a limiting density $g(x)$ and the sequence of knots $\underline{\tau}_K$ satisfies $\int_{\tau_{i-1}}^{\tau_i} p(t)dt = 1/K$, for a positive continuous density $p(t)$ on $[0,1]$, then the equivalent kernel for a general spline estimator satisfies

$$\mathcal{W}(x,t) = \frac{1}{g(t)h(k_q)/p(t)}\mathcal{K}\left\{\frac{x}{h(k_q)/p(t)}, \frac{t}{h(k_q)/p(t)}\right\}.$$

The question for equivalent kernel $\mathcal{K}(x,t)$ on bounded interval $[0,1]$ remains open. However, Theorem 2 emboldens to conjecture that it changes smoothly with k_q between $\mathcal{K}_{rs}^{[0,1]}(x,t)$ and $\mathcal{K}_{ss}^{[0,1]}(x,t)$ given in (3.34) and (3.17) correspondingly. Moreover, the boundary effects should also smoothly change with k_q. Indeed, for regression and smoothing splines, equivalent kernels on $[0,1]$ can be written as

$$\mathcal{K}_{rs}^{[0,1]}(x,t) = \mathcal{K}_{rs}(x,t) + \mathcal{K}_{rs}^{b}(x,t)$$
$$\mathcal{K}_{ss}^{[0,1]}(x,t) = \mathcal{K}_{ss}(x,t) + \mathcal{K}_{ss}^{b}(x,t),$$

where boundary kernels $\mathcal{K}_{rs}^{b}(x,t)$ and $\mathcal{K}_{ss}^{b}(x,t)$ are defined via

$$\mathcal{K}_{rs}^{b}(x,t) \coloneqq \mathcal{K}_{rs}^{[0,1]}(x,t) - \mathcal{K}_{rs}(x,t); \qquad \mathcal{K}_{ss}^{b}(x,t) \coloneqq \sum_{i=0}^{2q-1} \alpha_i(t)\psi_i(x),$$

with $\psi_i(x)$ given in (3.18) and $\alpha_i(t)$ arising from matching the natural boundary conditions. Kernel $\mathcal{K}_{rs}^{b}(x,t)$ is not available explicitly, while $\mathcal{K}_{ss}^{b}(x,t)$ has a complicated closed-form expression for each q. As we know from Huang and Studden (1992), at the boundaries, kernel $\mathcal{K}_{rs}^{[0,1]}(x,t)$ satisfies all conditions for boundary kernels as given in Gasser and Müller (1984), confirming that regression spline estimators do not have boundary effects. As scaled Green's function $\mathcal{R}_\lambda(x,t)$, kernel $\mathcal{K}_{ss}^{[0,1]}(x,t)$ satisfies natural boundary conditions (3.13) and causes the boundary effects of smoothing splines. Similarly, for penalized splines, we can define the equivalent kernel in interval $[0,1]$ as

$$\mathcal{K}^{[0,1]}(x,t) \coloneqq \mathcal{K}(x,t) + \mathcal{K}^{b}(x,t),$$

where $\mathcal{K}^{b}(x,t)$ is an unknown boundary kernel. As we know from Theorem 2, kernel $\mathcal{K}(x,t)$ varies smoothly between $\mathcal{K}_{rs}(x,t)$ and $\mathcal{K}_{ss}(x,t)$. Therefore, one can expect that additional boundary terms in $\mathcal{K}^{[0,1]}(x,t)$ to also vary smoothly between $\mathcal{K}_{rs}^{b}(x,t)$ and $\mathcal{K}_{ss}^{b}(x,t)$, so that the boundary effects of spline estimators grow as $k_q \to \infty$.

Extending our results onto d-dimensional space and studying the bias at the boundary in the spirit of Rice and Rosenblatt (1983) are interesting directions of further research.

Bibliography

Abramowitz, M. and I. Stegun (1972). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* New York: Dover.

Agarwal, G. G. and W. J. Studden (1980). Asymptotic integrated mean square error using least squares and bias minimizing splines. *The Annals of Statistics 8*(6), 1307–1325.

Barrow, D. L. and P. W. Smith (1978). Asymptotic properties of best $l_2[0,1]$ approximation by splines with variable knots. *Quart. Appl. Math. 36*, 293–304.

Berlinet, A. and C. Thomas-Agnan (2004). *Reproducing kernel Hilbert spaces in probability and statistics.* Kluwer Academic Publishers Boston.

Birkhoff, G. D. (1908). Boundary value and expansion problems of ordinary linear differential equations. *Trans. Amer. Math. Soc. 9*, 373–395.

Blu, T. and M. Unser (1999). Quantitative fourier analysis of approximation techniques: Part i – interpolators and projectors. *IEEE Transactions on signal processing 47*(10).

Claeskens, G., T. Krivobokova, and J. D. Opsomer (2009). Asymptotic properties of penalized spline estimators. *Biometrika 20*, 1–29.

Cogburn, R. and H. T. Davis (1974). Periodic splines and spectral estimations. *The Annals of Statistics 2*(6), 1108–1126.

Cox, D. D. (1983). Asymtotics for m-type smoothing splines. *The Annals of Statistics 11*(2), 530–551.

Cox, D. D. (1988). Approximation of method of regularization estimators. *The Annals of Statistics 16*(2), 694–712.

Craven, P. and G. Wahba (1979). Smoothing noisy data with spline functions. *Numerische Mathematik 31*(4), 377–403.

de Boor, C. (1978). *A Practical Guide to Splines.* Berlin: Springer.

Demmler, A. and C. Reinsch (1975). Oscillation matrices with spline smoothing. *Numer. Math. 24*(5), 375–382.

Eggermont, P. and V. LaRiccia (2006). Equivalent kernels for smoothing splines. *18*(2), 197–225.

Eggermont, P. and V. LaRiccia (2009). *Maximum Penalized Likelihood Estimation*, Volume 2. Springer Series in Statistics.

Gasser, T. and H. Müller (1979). Kernel estimation of regressions functions. *Smoothing Techniques for Curve Estimation*, 23–68.

Gasser, T. and H. Müller (1984). Estimation regression functions and their derivatives by the kernel method. *Ann. Statist. 11*, 210–229.

Gautschi, W. (1971). Attenuation factors in practical Fourier analysis. *Numerische Mathematik 18*(5), 373–400.

Golubev, G. and M. Nussbaum (1990). A risk bound in sobolev class regression. *The Annals of Statistics 18*(2), 758–778.

Hartley, H. (1961). The modified Gauss-Newton method for the fitting of nonlinear regression functions by least squares. *Technometrics 3*, 269–280.

He, T.-X. (2011). Eulerian polynomials and b-splines. *Journal of Computational and Applied Mathematics*.

Huang, S. and W. Studden (1990). Density estimation using spline projection kernels. Technical Report 90-52, Department of Statistics Purdue University.

Huang, S. and W. Studden (1992). An equivalent kernel method for least squares spline regression. Technical report, Department of Statistics Purdue University.

Huang, S. and W. Studden (1993). An equivalent kernel method for least squares spline regression. *Stat. Decis. 3*, 179–201.

Hudson, D. (1966). Fitting segmented curves whose join points have to be estimated. *J. Amer. Statist. Assoc. 61*, 1097–1129.

Kauermann, G., T. Krivobokova, and L. Fahrmeir (2009). Some asymptotic results on generalized penalized spline smoothing. *J. R. Stat. Soc. Ser. B Stat. Methodol. 71*(2), 487–503.

Lee, S., R. Tan, and W. Tang (1992). L_2-approximation by the translates of a function and related attenuation factors. *Numer. Math. 60*, 549–568.

Lee, S. and W. Tang (1991). Approximation and spectral properties of periodic spline operators. *Proceedings of the Edinburgh Mathematica Society 34*, 363–382.

Li, Y. and D. Ruppert (2008). On the asymptotics of penalized splines. *Biometrika 95*, 415–436.

Messer, K. (1991). A comparison of a spline estimate to its equivalent kernel estimate. *The Annals of Statistics 19*(2), 817–829.

Messer, K. and L. Goldstein (1993). A new class of kernels for nonparametric curve estimation. *The Annals of Statistics 21*(1), 179–195.

Nychka, D. (1995). Splines as local smoothers. *The Annals of Statistics 23*(4), 1175–1197.

Reinsch, C. (1967). Smoothing by spline functions. *Numer. Math. 10*, 177–183.

Rice, J. and M. Rosenblatt (1981). Integrated mean squared error of a smoothing spline. *Journal of approximation theory 33*, 353–369.

Rice, J. and M. Rosenblatt (1983). Smoothing splines: regression, derivatives and deconvolution. *The Annals of Statistics 11*(1), 141–156.

Ruppert, D., M. Wand, and R. Carroll (2003). *Semiparametric Regression.* Cambridge, UK: Cambridge University Press.

Schoenberg, I. (1964). Spline functions and the problem of graduation. *Proc. Nat. Acad. Sci. 52*, 947–950.

Schoenberg, I. (1973). *Cardinal spline interpolation.* Philadelphia: SIAM.

Schumaker, L. L. (Ed.) (2007). *Spline functions : basic theory.* Cambridge.

Silverman, B. W. (1985). Some aspects of the spline smoothing approach to non-parametric regression curve fitting. *47*, 1–21.

Speckman, P. (1985). Spline smoothing and optimal rates of convergence in nonparametric regression models. *The Annals of Statistics 13*(3), 970–983.

Stakgold, I. (1979). *Green's functions and boundary value problem.* A Wiley-Interscience publication.

Stone, C. J. (1980). Optimal rate of convergence for nonparametric estimators. *The Annals of Statistics 8*(6), 1348–1360.

Stone, C. J. (1982). Optimal global rates of convergence for nonparametric regression. *The Annals of Statistics 10*(4), 1040–1053.

Thomas-Agnan, C. (1996). Computing a family of reproducing kernels for statistical applications. *Numerical Algorithms 13*, 21–32.

Tsybakov, A. B. (2009). *Introduction to Nonparametric Estimation.* Springer Series in Statistics.

Utreras, F. (1980). Sur le choix du paramètre d'ajustement dans le lissage par fonctions spline. *Numer. Math. 34*, 15–28.

Utreras, F. (1983). Natural spline functions, their associated eigenvalue problem. *Numer. Math. 42*(1), 107–117.

Utreras, F. (1988). Boundary effects on convergence rates for tikhonov regularization. *Journal of approximation theory* (54), 235–249.

Wahba, G. (1975). Smoothing noisy data with spline functions. *Numer. Math. 24*, 383–393.

Wang, X., J. Shen, and D. Ruppert (2011). On the asymptotics of penalized spline smoothing. *Electronic Journal of Statistics 5*.

Whittaker, E. (1923). On a new method of graduation. *Proceedings of the Edinburgh Mathematica Society 41*, 63–74.

Xiao, L., Y. Li, T. Apanasovich, and D. Ruppert (2011). Local asymptotics of p-splines.

Zheludev, V. A. (1996). *Signal and Image Representation in Combined Spaces*, Chapter Periodic Splines, Harmonic Analysis and Wavelets, pp. 1–30. Josh Zeevi and Ronald Coifman.

Zhou, S., X. Shen, and D. A. Wolfe (1998). Local asymptotics for regression splines and confidence regions. *The Annals of Statistics 26*(5), 1760–1782.

Katsiaryna Schwarz

Place of birth: Minsk, Belarus

Nationalty: Belarusian

Education

10/2008-1/2013 **University of Göttingen, Germany**

PhD Student, Research Training Group 1023

10/2007-3/2008 **Friedrich Schiller University Jena, Germany**

Exchange student, School of Mathematics and Computer Science

9/2003-6/2008 **Belarussian State University, Minsk, Belarus**

Diploma in mathematics, Faculty of Mechanics and Mathematics

9/2001-6/2003 **Lyceum of Belarusian State University, Minsk, Belarus**

Mathematical group